Data-based systems and control theory

Day 1

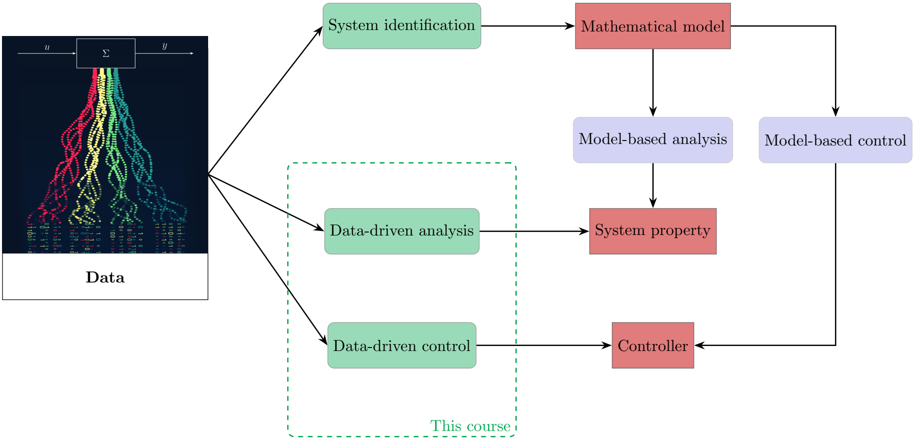
GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and **Henk van Waarde**

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen



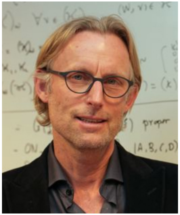
Motivation:

- We live in the era of big data
- Engineering systems are becoming more complex
- Direct approaches are promising in situations where modeling is challenging or computationally expensive

The course consists of five days:

1. Historical perspective
2. Data informativity, and analysis/design using noise-free data
3. Data-driven stabilization using noisy input-state data
4. Advanced topics: H_∞ control, dissipativity, and noisy input-output data
5. System identification and experiment design

1. Lecture: historical perspective
   - subspace identification
   - fundamental lemma
   - data-driven tracking
   - data-enabled predictive control
   - data-based closed-loop parameterization
2. Afternoon: problem solving session



Harry Trentelman

Jaap Eising

Paolo Rapisarda

Data informativity: Harry and Jaap
System identification: Paolo

Historical perspective

Subspace identification

Fundamental lemma

Data-driven tracking

Data-enabled predictive control

Data-based closed-loop parameterization

1985
1986
1989
2005
2008
2019
2020

Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Let $i, j \in \mathbb{N} = \{0, 1, \dots\}$ with $i \leqslant j$. Define $[i, j] := \{i, i+1, \dots, j\}$.
The restriction of $f : \mathbb{N} \to \mathbb{R}^q$ to $[i, j]$ is given by

$$f_{[i,j]} := \begin{bmatrix} f(i) \\ f(i+1) \\ \vdots \\ f(j) \end{bmatrix} \in \mathbb{R}^{(j-i+1)q}.$$

Let k be a positive integer such that $k \leqslant j - i + 1$. Define the **Hankel matrix** of depth k, associated with $f_{[i,j]}$, as

$$\mathcal{H}_k(f_{[i,j]}) := \begin{bmatrix} f(i) & f(i+1) & \cdots & f(j-k+1) \\ f(i+1) & f(i+2) & \cdots & f(j-k+2) \\ \vdots & \vdots & & \vdots \\ f(i+k-1) & f(i+k) & \cdots & f(j) \end{bmatrix} \in \mathbb{R}^{kq \times (j-i-k+2)}.$$

Definition: $f_{[i,j]}$ is said to be **persistently exciting** of order k if $\mathcal{H}_k(f_{[i,j]})$ has full row rank (equal to kq).

Consider the discrete-time **linear time-invariant system**

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{LinSys}$$

where $t \in \mathbb{N}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$.

Definition: We say (LinSys) is **(Kalman) controllable** if for every $x_0, x_1 \in \mathbb{R}^n$ there exist an $N \in \mathbb{N}$ and a sequence of inputs $u(0), u(1), \ldots, u(N-1) \in \mathbb{R}^m$ such that $x(N) = x_1$ whenever $x(0) = x_0$.

Kalman rank condition: The system (LinSys) is controllable **if and only if**

$$\operatorname{rank} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = n.$$

Terminology: The pair (A, B) is controllable.

The **(input-output) behavior** of (LinSys) is defined as:

$$\mathfrak{B} := \left\{ (u, y) : \mathbb{N} \to \mathbb{R}^{m+p} \mid \exists\, x : \mathbb{N} \to \mathbb{R}^n \text{ s.t. (LinSys) holds } \forall t \in \mathbb{N} \right\}.$$

Consider the discrete-time **linear time-invariant system**

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t). \end{aligned} \qquad \text{(LinSys)}$$

The **(input-output) behavior** of (LinSys) is defined as:

$$\mathfrak{B} := \left\{ (u, y) : \mathbb{N} \to \mathbb{R}^{m+p} \mid \exists\, x : \mathbb{N} \to \mathbb{R}^n \text{ s.t. (LinSys) holds } \forall t \in \mathbb{N} \right\}.$$

Definition: System (LinSys) is **(Kalman) observable** if for every $(u, y) \in \mathfrak{B}$ there is precisely one $x : \mathbb{N} \to \mathbb{R}^n$ such that (LinSys) holds for all $t \in \mathbb{N}$.

Terminology: The pair (C, A) is observable.

Kalman rank condition: The system (LinSys) is observable **if and only if**

$$\operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n.$$

Given $t_1, t_2 \in \mathbb{N}$ with $t_2 \geqslant t_1$, we define the **restricted behavior** as

$$\mathfrak{B}_{[t_1,t_2]} := \left\{ \begin{bmatrix} u_{[t_1,t_2]} \\ y_{[t_1,t_2]} \end{bmatrix} \mid (u,y) \in \mathfrak{B} \right\}.$$

Facts:

1. $\mathfrak{B}_{[t_1,t_2]}$ is a subspace of $\mathbb{R}^{(m+p)(t_2-t_1+1)}$
2. By time-invariance, $\mathfrak{B}_{[t_1,t_2]} \supseteq \mathfrak{B}_{[t_1+1,t_2+1]} \supseteq \mathfrak{B}_{[t_1+2,t_2+2]} \supseteq \cdots$
3. In general, equality does **not** hold
4. If (A,B) is **controllable** then $\mathfrak{B}_{[t_1,t_2]} = \mathfrak{B}_{[t_1+1,t_2+1]} = \mathfrak{B}_{[t_1+2,t_2+2]} = \cdots$

1985

1986 — Jan Willems. *From time series to linear system part II: exact modeling.*

1989 — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle. *On- and offline identification of linear state-space models*.

2005 — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor. *A note on persistency of excitation.*

2008 — Ivan Markovsky and Paolo Rapisarda. *Data-driven simulation and control.*

2019 — Jeremy Coulson, John Lygeros and Florian Dörfler. *Data-enabled predictive control: in the shallows of the DeePC.*

2020 — Claudio De Persis and Pietro Tesi. *Formulas for data-driven control: stabilization, optimality and robustness.*

2020 — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel. *Data informativity: a new perspective on data-driven analysis and control.*

2020 — From 2020 onwards: a very quickly growing literature.

Consider (LinSys) and assume that (A, B) is controllable and (C, A) is observable.

Problem: Given

$$\textbf{the data } \begin{bmatrix} u_{[0,T-1]} \\ y_{[0,T-1]} \end{bmatrix} \in \mathfrak{B}_{[0,T-1]} \text{ and an } \textbf{upper bound } N \geqslant n,$$

find the state-space dimension n and **matrices** $\bar{A}, \bar{B}, \bar{C}$ and $\bar{D}$ such that

$$\bar{A} = SAS^{-1}, \ \bar{B} = SB, \ \bar{C} = CS^{-1}, \quad \text{and} \ \bar{D} = D$$

for some nonsingular matrix $S \in \mathbb{R}^{n \times n}$.

So the goal is to **identify the system** up to a similarity transformation!

Thought experiment: Suppose that also $x_{[0,T]}$ is given.
Then (A, B, C, D) is a solution to the system of linear equations:

$$\begin{bmatrix} x(1) & x(2) & \cdots & x(T) \\ y(0) & y(1) & \cdots & y(T-1) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \\ u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}.$$

The approach of Moonen et al.:
Assume $T \geqslant 2N$ and consider the partitioned Hankel matrix

$$\mathcal{H}_{2N}(u_{[0,T-1]}) = \left[\begin{array}{cccc} u(0) & u(1) & \cdots & u(T-2N) \\ \vdots & \vdots & & \vdots \\ u(N-1) & u(N) & \cdots & u(T-N-1) \\ \hdashline u(N) & u(N+1) & \cdots & u(T-N) \\ \vdots & \vdots & & \vdots \\ u(2N-1) & u(2N) & \cdots & u(T-1) \end{array}\right] =: \left[\begin{array}{c} U_p \\ \hdashline U_f \end{array}\right].$$

Similarly, partition $\mathcal{H}_{2N}(y_{[0,T-1]})$ into the blocks Y_p and Y_f.

Define the state matrices (**not given!**):

$$\begin{aligned} X_p &:= \begin{bmatrix} x(0) & x(1) & \cdots & x(T-2N) \end{bmatrix} \\ X_f &:= \begin{bmatrix} x(N) & x(N+1) & \cdots & x(T-N) \end{bmatrix}. \end{aligned}$$

Theorem: If the matrix

$$\begin{bmatrix} X_p \\ \hdashline U_p \\ U_f \end{bmatrix} = \begin{bmatrix} \mathcal{H}_1(x_{[0,T-2N]}) \\ \hdashline \mathcal{H}_{2N}(u_{[0,T-1]}) \end{bmatrix}$$

has rank $n + 2mN$ then $\operatorname{rsp} X_f = \operatorname{rsp} \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \cap \operatorname{rsp} \begin{bmatrix} U_f \\ Y_f \end{bmatrix}$.

So a **state sequence** can be obtained **from input-output data**!

Remark: Idea of state reconstruction (using an infinite Hankel matrix) used already in 1986 by Willems (From times series paper part II).

An issue: The rank condition is **not verifiable from data**...

A (coarse) timeline

1985

- 1986 — Jan Willems.
  From time series to linear system part II: exact modeling.
- 1989 — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
  On- and offline identification of linear state-space models.
- 2005 — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
  A note on persistency of excitation.
- 2008 — Ivan Markovsky and Paolo Rapisarda.
  Data-driven simulation and control.
- 2019 — Jeremy Coulson, John Lygeros and Florian Dörfler.
  Data-enabled predictive control: in the shallows of the DeePC.
- 2020 — Claudio De Persis and Pietro Tesi.
  Formulas for data-driven control: stabilization, optimality and robustness.
- 2020 — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
  Data informativity: a new perspective on data-driven analysis and control.
- 2020 — From 2020 onwards: a very quickly growing literature.

Again, consider the **data**: $\begin{bmatrix} u_{[0,T-1]} \\ y_{[0,T-1]} \end{bmatrix} \in \mathfrak{B}_{[0,T-1]}$.

Let $L \in [1, T]$. Each column of

$$\begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \hdashline \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-L) \\ \vdots & \vdots & & \vdots \\ u(L-1) & u(L) & \cdots & u(T-1) \\ \hdashline y(0) & y(1) & \cdots & y(T-L) \\ \vdots & \vdots & & \vdots \\ y(L-1) & y(L) & \cdots & y(T-1) \end{bmatrix} \tag{1}$$

is in the restricted behavior $\mathfrak{B}_{[0,L-1]}$.

Since $\mathfrak{B}_{[0,L-1]}$ is a subspace, **all linear combinations of the columns of (1)** are also in $\mathfrak{B}_{[0,L-1]}$.

So all linear combinations of the columns of (1) are in $\mathfrak{B}_{[0,L-1]}$, in other words:

$$\operatorname{im} \begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} \subseteq \mathfrak{B}_{[0,L-1]}.$$

Important question: Under which conditions do we have

$$\operatorname{im} \begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} = \mathfrak{B}_{[0,L-1]} \text{ ?}$$

This would allow us to **parameterize** all length-L trajectories using **data**:

$$\begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]} \iff \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} = \begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} g$$

for some $g \in \mathbb{R}^{T-L+1}$.

Theorem: Let $L \in [1, T]$. Assume that (A, B) is **controllable** and $u_{[0,T-1]}$ is **persistently exciting** of order $n + L$. Then:

1. $\operatorname{rank} \begin{bmatrix} \mathcal{H}_1(x_{[0,T-L]}) \\ \mathcal{H}_L(u_{[0,T-1]}) \end{bmatrix} = n + mL.$
2. $\operatorname{im} \begin{bmatrix} \mathcal{H}_L(u_{[0,T-1]}) \\ \mathcal{H}_L(y_{[0,T-1]}) \end{bmatrix} = \mathfrak{B}_{[0,L-1]}.$

Remarks:

- We can thus **parameterize all length-L trajectories** using data!
- PE of order $n + L$ requires $T \geqslant (n + L)(m + 1) - 1$
- If only upper bound $N \geqslant n$ is given, use $u_{[0,T-1]}$ that is **PE of order $N + L$**
- If $L > \ell$ where ℓ is the smallest integer such that

  $$\operatorname{rank}\begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{\ell-1})^\top \end{bmatrix} = \operatorname{rank}\begin{bmatrix} C^\top & (CA)^\top & \cdots & (CA^{\ell})^\top \end{bmatrix}$$

  (i.e., the **lag** of the system), then $\mathfrak{B}_{[0,L-1]}$ uniquely determines $\mathfrak{B}$
- Rank conditions like 1 are relevant for **subspace identification** $(L = 2N)$

Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. The **reachable subspace** is defined as:

$$\langle A \mid \operatorname{im} B\rangle := \operatorname{im} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$$

So (A, B) is controllable if and only if $\langle A \mid \operatorname{im} B\rangle = \mathbb{R}^n$.

Definition: A subspace $\mathcal{V} \subseteq \mathbb{R}^n$ is called **A-invariant** if $A\mathcal{V} \subseteq \mathcal{V}$.

Facts:

1. $\langle A \mid \operatorname{im} B\rangle$ is A-invariant and contains $\operatorname{im} B$.
2. $\langle A \mid \operatorname{im} B\rangle$ is the **smallest** subspace with the above two properties. That is, if $\mathcal{V} \subseteq \mathbb{R}^n$ is A-invariant and contains $\operatorname{im} B$ then $\langle A \mid \operatorname{im} B\rangle \subseteq \mathcal{V}$.

Lemma: Let $k \geqslant 1$ be an integer. Then (A, B) is controllable **if and only if** $(\mathcal{A}_k, \mathcal{B}_k)$ is controllable, where

$$\mathcal{A}_k := \begin{bmatrix} A & B & 0 \\ 0 & 0 & I_{(k-1)m} \\ 0_{m\times n} & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n+km)\times(n+km)}, \quad \mathcal{B}_k := \begin{bmatrix} 0 \\ 0 \\ I_m \end{bmatrix} \in \mathbb{R}^{(n+km)\times m}.$$

Claim: Define the matrix

$$J_k := \begin{bmatrix} \mathcal{H}_1(x_{[0,T-k]}) \\ \mathcal{H}_k(u_{[0,T-1]}) \end{bmatrix} \qquad \text{for } k = 1, 2, \ldots, n+L.$$

Assume that $\operatorname{rank} J_q < n + mq$ for some $q \in \{1, 2, \ldots, n+L\}$. Then

$$\operatorname{rank} J_k < n + mk$$

for all $k = q, q+1, \ldots, n+L$.

Proof of the claim on the board

Proof of fundamental lemma – statement 1:

- We want to show that $\operatorname{rank} J_L = n + mL$.
- Suppose on the contrary that $\operatorname{rank} J_L < n + mL$.
- By the **claim**, $\operatorname{rank} J_k < n + mk$ **for all** $k = L, L+1, \ldots, n+L$.
- **Proof by contraction on the board.** $\square$

1985
1986
1989
2005
2008
2019
2020

Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Historical perspective

Subspace identification

Fundamental lemma

Data-driven tracking

Data-enabled predictive control

Data-based closed-loop parameterization

Problem: Let $L_{\text{ini}}, L_r \in \mathbb{N}$ be positive integers and define $L = L_{\text{ini}} + L_r$. Given:

- a symmetric positive semidefinite **(weight) matrix**

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \in \mathbb{R}^{(m+p)\times(m+p)},$$

- the **data** $\begin{bmatrix} u_{[0,T-1]} \\ y_{[0,T-1]} \end{bmatrix} \in \mathfrak{B}_{[0,T-1]}$,
- an **initial trajectory** $\begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L_{\text{ini}}-1]}$,
- and a **reference signal** $(v_{[L_{\text{ini}},L-1]}, z_{[L_{\text{ini}},L-1]}) \in \mathbb{R}^{mL_r} \times \mathbb{R}^{pL_r}$,

find a trajectory $(\bar{u}_{[L_{\text{ini}},L-1]}, \bar{y}_{[L_{\text{ini}},L-1]})$ that solves

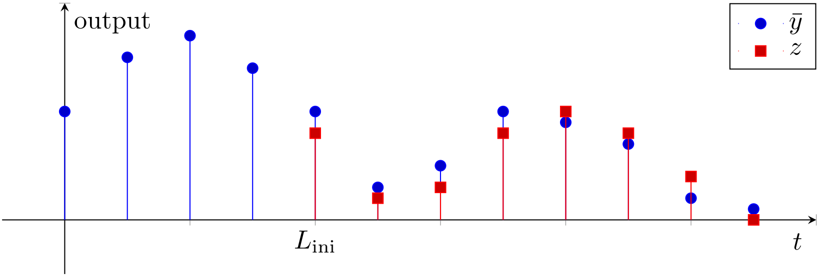
$$\text{minimize} \sum_{t=L_{\text{ini}}}^{L-1} \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}^\top Q \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}$$

$$\text{subject to} \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]}.$$

Problem: find a trajectory $(\bar{u}_{[L_{\mathsf{ini}},L-1]}, \bar{y}_{[L_{\mathsf{ini}},L-1]})$ that solves

$$\text{minimize} \quad \sum_{t=L_{\mathsf{ini}}}^{L-1} \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}^\top Q \begin{bmatrix} \bar{u}(t) - v(t) \\ \bar{y}(t) - z(t) \end{bmatrix}$$

$$\text{subject to} \quad \begin{bmatrix} \bar{u}_{[0,L-1]} \\ \bar{y}_{[0,L-1]} \end{bmatrix} \in \mathfrak{B}_{[0,L-1]}.$$



Notation: write $L = L_{\text{ini}} + L_r$ and partition

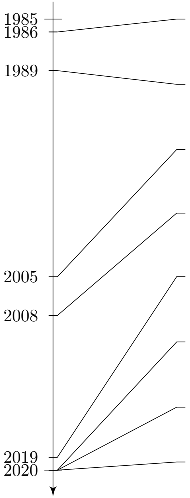
$$\mathcal{H}_L(u_{[0,T-1]}) = \begin{bmatrix} U_p \\ U_f \end{bmatrix} \quad \text{and} \quad \mathcal{H}_L(y_{[0,T-1]}) = \begin{bmatrix} Y_p \\ Y_f \end{bmatrix}$$

Under the assumptions of the **fundamental lemma**, the **constraint is replaced** by:

$$\begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \\ \bar{u}_{[L_{\text{ini}},L-1]} \\ \bar{y}_{[L_{\text{ini}},L-1]} \end{bmatrix} = \begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix} g \quad \text{for some} \quad g \in \mathbb{R}^{T-L+1}.$$

By defining $\tilde{Q} := \begin{bmatrix} I_{L_r} \otimes Q_{11} & I_{L_r} \otimes Q_{12} \\ I_{L_r} \otimes Q_{21} & I_{L_r} \otimes Q_{22} \end{bmatrix}$, the problem boils down to:

$$\begin{aligned} \text{minimize} \quad & \left\| \tilde{Q}^{\frac{1}{2}} \begin{bmatrix} U_f \\ Y_f \end{bmatrix} g - \tilde{Q}^{\frac{1}{2}} \begin{bmatrix} v_{[L_{\text{ini}},L-1]} \\ z_{[L_{\text{ini}},L-1]} \end{bmatrix} \right\|^2 \\ \text{subject to} \quad & \begin{bmatrix} U_p \\ Y_p \end{bmatrix} g = \begin{bmatrix} \bar{u}_{[0,L_{\text{ini}}-1]} \\ \bar{y}_{[0,L_{\text{ini}}-1]} \end{bmatrix}. \end{aligned} \qquad \text{(constrained LS)}$$



Jan Willems.
From time series to linear system part II: exact modeling.

Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
On- and offline identification of linear state-space models.

Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
A note on persistency of excitation.

Ivan Markovsky and Paolo Rapisarda.
Data-driven simulation and control.

Jeremy Coulson, John Lygeros and Florian Dörfler.
Data-enabled predictive control: in the shallows of the DeePC.

Claudio De Persis and Pietro Tesi.
Formulas for data-driven control: stabilization, optimality and robustness.

Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
Data informativity: a new perspective on data-driven analysis and control.

From 2020 onwards: a very quickly growing literature.

Historical perspective

Subspace identification

Fundamental lemma

Data-driven tracking

Data-enabled predictive control

Data-based closed-loop parameterization

Given: data, Q, initial trajectory, and **reference signal** $\{(v(t), z(t))\}_{t=L_{\text{ini}}}^{\infty}$.

1. Set $\tau = 0$
2. Compute a trajectory $(u^{\text{pred}}_{[\tau+L_{\text{ini}},\tau+L-1]}, y^{\text{pred}}_{[\tau+L_{\text{ini}},\tau+L-1]})$ solving

$$\text{minimize} \quad \sum_{t=\tau+L_{\text{ini}}}^{\tau+L-1} \begin{bmatrix} u^{\text{pred}}(t) - v(t) \\ y^{\text{pred}}(t) - z(t) \end{bmatrix}^\top Q \begin{bmatrix} u^{\text{pred}}(t) - v(t) \\ y^{\text{pred}}(t) - z(t) \end{bmatrix}$$

$$\text{subject to} \quad \begin{bmatrix} \bar{u}_{[\tau,\tau+L_{\text{ini}}-1]} \\ u^{\text{pred}}_{[\tau+L_{\text{ini}},\tau+L-1]} \\ \bar{y}_{[\tau,\tau+L_{\text{ini}}-1]} \\ y^{\text{pred}}_{[\tau+L_{\text{ini}},\tau+L-1]} \end{bmatrix} \in \mathfrak{B}_{[\tau,\tau+L-1]}.$$

3. Apply $\bar{u}(\tau + L_{\text{ini}}) = u^{\text{pred}}(\tau + L_{\text{ini}})$ to the system, and measure $\bar{y}(\tau + L_{\text{ini}})$
4. Set $\tau \leftarrow \tau + 1$ and go to step 2.

Remarks on generalizations:

- Equivalence with **model predictive control (MPC)**
- **Constraints** on inputs and outputs, regularization methods
- **Stability analysis** under terminal constraints[1]

[1] Berberich et al., *Data-driven model predictive control with stability and robustness guarantees*, IEEE TAC, 2020

- 1985
- 1986 — Jan Willems. *From time series to linear system part II: exact modeling.*
- 1989 — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle. *On- and offline identification of linear state-space models.*
- 2005 — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor. *A note on persistency of excitation.*
- 2008 — Ivan Markovsky and Paolo Rapisarda. *Data-driven simulation and control.*
- 2019 — Jeremy Coulson, John Lygeros and Florian Dörfler. *Data-enabled predictive control: in the shallows of the DeePC.*
- 2020 — Claudio De Persis and Pietro Tesi. *Formulas for data-driven control: stabilization, optimality and robustness.*
- 2020 — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel. *Data informativity: a new perspective on data-driven analysis and control.*
- 2020 — From 2020 onwards: a very quickly growing literature.

Historical perspective

Subspace identification

Fundamental lemma

Data-driven tracking

Data-enabled predictive control

Data-based closed-loop parameterization

Now consider the **input-state dynamics** $x(t+1) = Ax(t) + Bu(t)$, and the **input-state** data $X := \mathcal{H}_1(x_{[0,T]})$ and $U_- := \mathcal{H}_1(u_{[0,T-1]})$.

Define $X_+ := \mathcal{H}_1(x_{[1,T]})$ and $X_- := \mathcal{H}_1(x_{[0,T-1]})$. **Note**: $X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}$.

Assume that (A, B) is **controllable** and $u_{[0,T-1]}$ is **PE of order** $n+1$. Then

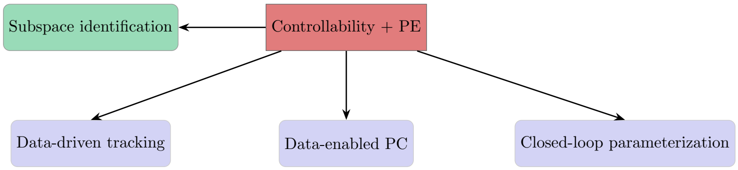
$$\operatorname{rank} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m.$$

Thus, for any feedback controller $u = Kx$, there exists $G \in \mathbb{R}^{T \times n}$ such that

$$\begin{bmatrix} I \\ K \end{bmatrix} = \begin{bmatrix} X_- \\ U_- \end{bmatrix} G.$$

This leads to the **data-based representation** of the closed-loop system:

$$x(t+1) = (A + BK)x(t) = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} Gx(t) = X_+ Gx(t).$$



Important questions:

1. So far, all methods work with **persistently exciting input data**
   - can we find controllers in situations where the input is **not PE**?
   - can we do so in situations where **unique identification is not possible**?
2. what is the **minimum number of samples** needed for control design?
3. and what about **noisy** data?

We will need a **new framework** to study these questions...

1985

- 1986 — Jan Willems.
  From time series to linear system part II: exact modeling.
- 1989 — Marc Moonen, Bart De Moor, Lieven Vandenberghe and Joos Vandewalle.
  On- and offline identification of linear state-space models.
- 2005 — Jan Willems, Paolo Rapisarda, Ivan Markovsky and Bart De Moor.
  A note on persistency of excitation.
- 2008 — Ivan Markovsky and Paolo Rapisarda.
  Data-driven simulation and control.
- 2019 — Jeremy Coulson, John Lygeros and Florian Dörfler.
  Data-enabled predictive control: in the shallows of the DeePC.
- 2020 — Claudio De Persis and Pietro Tesi.
  Formulas for data-driven control: stabilization, optimality and robustness.
- 2020 — Henk van Waarde, Jaap Eising, Harry Trentelman and Kanat Camlibel.
  Data informativity: a new perspective on data-driven analysis and control.
- 2020 — From 2020 onwards: a very quickly growing literature.

1. **Historical perspective**
   - subspace identification
   - fundamental lemma
   - data-driven tracking
   - data-enabled predictive control
   - data-based closed-loop parameterization
2. Tomorrow, we will introduce the **data informativity framework** to
   - deal with data that are **not persistently exciting**
   - find controllers in situations where system identification is **not possible**
   - deal with **noisy measurements**

Data-based systems and control theory

Day 2

GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and Henk van Waarde

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen

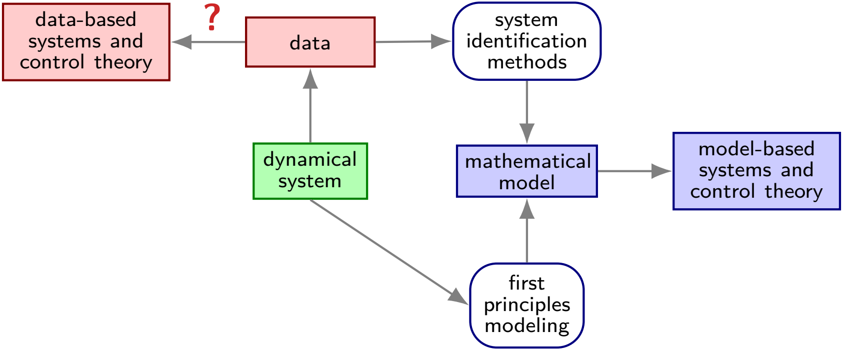
Model-based vs. data-based systems and control theory

Data informativity framework

Data informativity with noise-free data (analysis)

Data informativity with noise-free data (design)

Data informativity with noisy data


data-based
systems and
control theory
?
data
system
identification
methods
dynamical
system
mathematical
model
model-based
systems and
control theory
first
principles
modeling

$\mathcal{M}$ model class (non)linear state-space models, transfer matrix models, etc.

$\mathcal{S} \in \mathcal{M}$ **given** model

analysis: given a property $\mathcal{P}$ stability, controllability, etc.

verify if $\mathcal{S}$ has the property $\mathcal{P}$

design: given a control objective $\mathcal{O}$ stabilization, optimal control, etc.

find a controller $\mathcal{C}$ s.t. the closed-loop system $\mathcal{S} \wedge \mathcal{C}$ achieves the objective $\mathcal{O}$

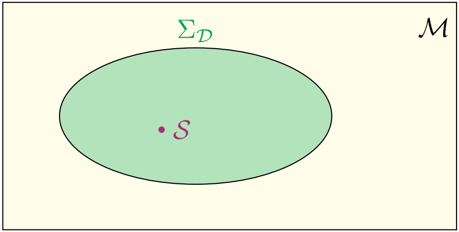
Model-based vs. data-based systems and control theory

Data informativity framework

Data informativity with noise-free data (analysis)

Data informativity with noise-free data (design)

Data informativity with noisy data



$\mathcal{M}$ model class
$\mathcal{S}$ unknown system
$\mathcal{D}$ given data
$\Sigma_{\mathcal{D}}$ set of data-consistent systems

$\mathcal{P}$ system property
$\mathcal{O}$ control objective

Data $\mathcal{D}$ are informative[1]

for $\mathcal{P}$ $:\Longleftrightarrow$ $\mathcal{P}$ holds **for all systems in** $\Sigma_{\mathcal{D}}$

for $\mathcal{O}$ $:\Longleftrightarrow$ $\exists$ controller $\mathcal{C}$ that achieves $\mathcal{O}$ **for all systems in** $\Sigma_{\mathcal{D}}$

[1] van Waarde, Eising, Trentelman, and Camlibel, "Data informativity: a new perspective on data-driven analysis and control", *IEEE TAC*, 2020.

Model-based vs. data-based systems and control theory

Data informativity framework

Data informativity with noise-free data (analysis)

Data informativity with noise-free data (design)

Data informativity with noisy data

- limited applicability as measurements are never exact (noise-free)
- reveals fundamental limitations of data-driven control
- a stepping stone for studying noisy data

Setup

state data

model class $\boldsymbol{x}(t+1) = A\boldsymbol{x}(t)$ $\quad \boldsymbol{x} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$

true system $\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t)$ $\quad A_s$ is unknown

the data collected from the true system

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix}$$

satisfies

$$X_+ = A_s X_-$$

where

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

data-consistent systems

$$\Sigma = \{A : X_+ = AX_-\} \qquad A_s \in \Sigma$$

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{A : X_+ = AX_-\} \qquad A_s \in \Sigma$$

Definition system identification

The data X are informative for system identification if $\Sigma = \{A_s\}$.

Theorem system identification

The data X are informative for system identification $\iff$ X_- has full row rank.

Proof

$X_+ = AX_-$ has a unique solution A $\iff$ X_- has full row rank.

Observation $(X_- X_-^\dagger = I_n)$

If X_- has full row rank, then it has **a** right inverse. For any right inverse $X_-^\dagger$, we have $X_+ X_-^\dagger = A_s X_- X_-^\dagger = A_s$.

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{A : X_+ = AX_-\} \qquad A_s \in \Sigma$$

Definition stability [A is stable $:\iff$ $|\lambda| < 1$ for all $\lambda \in \sigma(A)$]

The data X are informative for stability if A is stable for all $A \in \Sigma$.

Theorem stability

The data X are informative for stability $\iff$ X_- has full row rank and $X_+X_-^\dagger$ is stable.

Proof

- if: X_- has full row rank $\implies$ $\Sigma = \{A_s\}$ and $A_s = X_+X_-^\dagger$ is stable.
- only if: Let ξ be s.t. $\xi^\top X_- = 0$, $A \in \Sigma$, define $A_\alpha := A + \alpha\xi\xi^\top$. Then, $A_\alpha \in \Sigma$.

  $\implies$ A_α is stable for all α $\implies$ $n > |\operatorname{trace} A_\alpha| = |\operatorname{trace} A + \alpha\,\xi^\top\xi|$ for all α.

  $\implies$ $\xi = 0 \implies$ X_- has full row rank $\implies$ $X_+X_-^\dagger = A_s$ is stable.

Theorem system identification

The data X are informative for system identification $\iff$ X_- has full row rank.

Theorem stability

The data X are informative for stability $\iff$ X_- has full row rank and $X_+X_-^\dagger$ is stable.

Observation

One cannot verify stability without being able to identify the true system!

Setup

input-state data

model class $\boldsymbol{x}(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t)$ $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{u} \in \mathbb{R}^m$

true system $\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) + B_s\boldsymbol{u}(t)$ A_s and B_s are unknown

the data collected from the true system

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

satisfy

$$X_+ = A_s X_- + B_s U_-$$

where

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

data-consistent systems

$$\Sigma = \{(A, B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A, B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition system identification

The data (U_-, X) are informative for system identification if $\Sigma = \{(A_s, B_s)\}$.

Theorem system identification

The data (U_-, X) are informative for SI $\iff \begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has full row rank.

Proof

$X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has a unique solution $\begin{bmatrix} A & B \end{bmatrix} \iff \begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has full row rank.

Observation

If $\begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has full row rank, then $\begin{bmatrix} A_s & B_s \end{bmatrix} = X_+ \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\dagger$.

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A,B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Question

Under what conditions is the unknown system (A_s, B_s) controllable?

Observation sufficient condition

- $\begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has full row rank $\implies \begin{bmatrix} A_s & B_s \end{bmatrix} = X_+ \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\dagger$
- check if (A_s, B_s) is controllable!

Question necessity

Does $\begin{bmatrix} X_- \\ U_- \end{bmatrix}$ need to have full row rank to conclude controllability of (A_s, B_s)?

Recall — data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A, B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition — controllability

(U_-, X) are informative for controllability if (A, B) is controllable for all $(A, B) \in \Sigma$.

Recall — Hautus test

(A, B) is controllable $\iff$ $\begin{bmatrix} A - \lambda I & B \end{bmatrix}$ has full row rank for all $\lambda \in \sigma(A)$.

Theorem — controllability

(U_-, X) are informative for controllability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.

Definition controllability

(U_-, X) are informative for controllability if (A, B) is controllable for all $(A, B) \in \Sigma$.

Theorem controllability

(U_-, X) are informative for controllability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.

Example verifying controllability of an unknown system

Let $n = 2$ and $m = 1$. Consider the data

$$X = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Then, we have $X_- = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $X_+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies X_+ - \lambda X_- = \begin{bmatrix} 1 & -\lambda \\ 0 & 1 \end{bmatrix}$

$\overset{(1.22)}{\Longrightarrow}$ data are informative for controllability

Every system of the form $(A, B) = \left(\begin{bmatrix} 0 & a \\ 1 & b \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$ generates the data!

Controllability of (A_s, B_s) can be deduced from data although we **cannot identify** it!

Theorem controllability

(U_-, X) are informative for controllability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.

Proof

- if: Suppose that $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.
  Let $(A, B) \in \Sigma \implies X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \implies X_+ - \lambda X_- = \begin{bmatrix} A - \lambda I & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}$
  $\implies \begin{bmatrix} A - \lambda I & B \end{bmatrix}$ has full row rank for all $\lambda \in \mathbb{C} \implies (A, B)$ is controllable!
- only if: Suppose that (A, B) is controllable for all $(A, B) \in \Sigma$.
  goal: $\lambda \in \mathbb{C}$, $z \in \mathbb{C}^n$, $z^*(X_+ - \lambda X_-) = 0 \implies z = 0$
  - ◇ **case 1:** $\lambda \in \mathbb{R}$ and $z \in \mathbb{R}^n$ with $z^\top z = 1$
    Let $(A, B) \in \Sigma \implies z^\top (X_+ - \lambda X_-) = z^\top \begin{bmatrix} A - \lambda I & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0$
    $$\implies \begin{bmatrix} A - zz^\top (A - \lambda I) & B - zz^\top B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = X_+$$
    $\implies (A - zz^\top (A - \lambda I), B - zz^\top B) \in \Sigma$ and hence is controllable!
  $z^\top \left(A - zz^\top (A - \lambda I)\right) = \lambda z^\top$ and $z^\top \left(B - zz^\top B\right) = 0 \implies z = 0$ (contradiction!)
  - ◇ **case 2:** $\lambda \in \mathbb{C} \setminus \mathbb{R}$ and $z \in \mathbb{C}^n$ with $z^* z = 1$
    similar but a bit more involved argument

Definition controllability

(U_-, X) are informative for controllability if (A, B) is controllable for all $(A, B) \in \Sigma$.

Theorem controllability

(U_-, X) are informative for controllability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.

Recall Hautus test

(A, B) is controllable $\iff$ $\operatorname{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix}$ has full row rank for all $\lambda \in \sigma(A)$.

Theorem data-driven Hautus test

The data (U_-, X) are informative for controllability $\iff$

(a) X_+ has full row rank, and

(b) $\mu X_+ - X_-$ has full row rank for all $\mu \in \sigma(X_- X_+^\dagger)$.

Theorem controllability

(U_-, X) are informative for controllability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$.

Theorem data-driven Hautus test

The data (U_-, X) are informative for controllability $\iff$

(a) X_+ has full row rank, and

(b) $\mu X_+ - X_-$ has full row rank for all $\mu \in \sigma(X_- X_+^\dagger)$.

Definition stabilizability

(U_-, X) are informative for stabilizability if (A, B) is stabilizable for all $(A, B) \in \Sigma$.

Theorem stabilizability

(U_-, X) are informative for stabilizability $\iff$ $X_+ - \lambda X_-$ has full row rank for all $\lambda \in \mathbb{C}$ with $|\lambda| \geqslant 1$.

Theorem data-driven Hautus test

The data (U_-, X) are informative for stabilizability $\iff$

(a) $X_+ - X_-$ has full row rank, and

(b) $\mu(X_+ - X_-) - X_-$ has full row rank for all $\mu \in \sigma(X_-(X_+ - X_-)^\dagger)$.

Recall — data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A,B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition — stability

The data (U_-, X) are informative for stability if A is stable for all $(A,B) \in \Sigma$.

Theorem — stability

The data (U_-, X) are informative for stability $\iff$

(a) X_- has full row rank, and

(b) there exists a right inverse $X_-^\dagger$ such that $X_+X_-^\dagger$ is stable and $U_-X_-^\dagger = 0$.

In that case, we have $A = X_+X_-^\dagger$ for all $(A,B) \in \Sigma$, so in particular $A_s = X_+X_-^\dagger$.

Theorem — system identification

The data (U_-, X) are informative for SI $\iff$ $\begin{bmatrix} X_- \\ U_- \end{bmatrix}$ has full row rank.

Setup

state-output data

model class $\boldsymbol{x}(t+1) = A\boldsymbol{x}(t) \qquad \boldsymbol{y}(t) = C\boldsymbol{x}(t) \qquad \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^p$

true system $\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) \qquad \boldsymbol{y}(t) = C_s\boldsymbol{x}(t) \qquad A_s$ and C_s are unknown

the data collected from the true system

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad Y_- = \begin{bmatrix} y(0) & y(1) & \cdots & y(T-1) \end{bmatrix}$$

satisfy

$$\begin{bmatrix} X_+ \\ Y_- \end{bmatrix} = \begin{bmatrix} A_s \\ C_s \end{bmatrix} X_-.$$

where

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

data-consistent systems

$$\Sigma = \left\{(C, A) : \begin{bmatrix} X_+ \\ Y_- \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} X_- \right\} \qquad (C_s, A_s) \in \Sigma$$

Recall
data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad Y_- = \begin{bmatrix} y(0) & y(1) & \cdots & y(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \left\{ (C, A) : \begin{bmatrix} X_+ \\ Y_- \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} X_- \right\} \qquad (C_s, A_s) \in \Sigma$$

Definition
The data (X, Y_-) are informative for system identification if $\Sigma = \{(C_s, A_s)\}$.

Theorem
The data (X, Y_-) are informative for system identification $\iff$ X_- has full row rank.

Definition
(X, Y_-) are informative for observability if (C, A) is observable for all $(C, A) \in \Sigma$.

Theorem
The data (X, Y_-) are informative for observability $\iff$

(a) X_- has full row rank, and

(b) $(Y_- X_-^\dagger, X_+ X_-^\dagger)$ is observable.

model	data	informative for	conditions
$\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t)$	X	stability	X_- full row rank and $X_+X_-^\dagger$ is stable
$\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) + B_s\boldsymbol{u}(t)$	(U_-, X)	controllability	$X_+ - \lambda X_-$ full row rank for all $\lambda \in \mathbb{C}$
$\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) + B_s\boldsymbol{u}(t)$	(U_-, X)	stabilizability	$X_+ - \lambda X_-$ full row rank for all $\lambda \in \mathbb{C}$ with $\lvert\lambda\rvert \geqslant 1$
$\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) + B_s\boldsymbol{u}(t)$	(U_-, X)	stability	X_- full row rank, $X_+X_-^\dagger$ is stable and $U_-X_-^\dagger = 0$
$\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t)$ $\boldsymbol{y}(t) = C_s\boldsymbol{x}(t)$	(X, Y_-)	observability	X_- full row rank and $(YX_-^\dagger, X_+X_-^\dagger)$ is observable

Model-based vs. data-based systems and control theory

Data informativity framework

Data informativity with noise-free data (analysis)

Data informativity with noise-free data (design)

Data informativity with noisy data

Recall — data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A, B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition — stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback if there exists K such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Question — from stabilizability to stabilization by state feedback

Is informativity for stabilizability enough?

Example — stabilizability does not imply stabilization by state feedback

Let $n = m = 1$. Consider the data $X = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $U_- = 1$.

Clearly, $X_- = 0$ and $X_+ = 1$. Then, $X_+ - \lambda X_-$ has full row rank for all λ with $|\lambda| \geqslant 1$.

This means that (U_-, X) are informative for controllability and hence for stabilizability.

Note that $\Sigma = \{(a, 1) : a \in \mathbb{R}\}$. So, $(-1, 1), (1, 1) \in \Sigma$.

However, there is no k such that both $-1 + k$ and $1 + k$ are stable.

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A, B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback if there exists K such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Theorem stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback $\iff$

(a) X_- has full row rank, and

(b) there exists **a** right inverse $X_-^\dagger$ such that $X_+X_-^\dagger$ is stable.

Moreveor, if these conditions are satisfied, then $K := U_-X_-^\dagger$ is a stabilizing feedback.

Definition stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback if there exists K such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Theorem stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback $\iff$

(a) X_- has full row rank, and

(b) there exists **a** right inverse $X_-^\dagger$ such that $X_+ X_-^\dagger$ is stable.

Example system identification is **NOT** necessary for stabilization

Let $n = 2$ and $m = 1$. Consider $X = \begin{bmatrix} 1 & 0.5 & -0.25 \\ 0 & 1 & 1 \end{bmatrix}$ and $U_- = \begin{bmatrix} -1 & -1 \end{bmatrix}$.

$$\implies X_- = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} \text{ and } X_+ = \begin{bmatrix} 0.5 & -0.25 \\ 1 & 1 \end{bmatrix} \implies X_+ X_-^\dagger = \begin{bmatrix} 0.5 & -0.5 \\ 1 & 0.5 \end{bmatrix}$$

$\lambda_{1,2} = \frac{1}{2}(1 \pm \sqrt{2}i) \overset{(1.4)}{\implies}$ informativity for stabilization by state feedback.

$(A, B) = \left(\begin{bmatrix} 1.5 + a & 0.5a \\ 1 + b & 0.5 + 0.5b \end{bmatrix}, \begin{bmatrix} 1 + a \\ b \end{bmatrix} \right)$ generates the data for every a, b!

The feedback $K = \begin{bmatrix} -1 & -0.5 \end{bmatrix}$ stabilizes all data-consistent systems!

Theorem stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback $\iff$

(a) X_- has full row rank, and

(b) there exists **a** right inverse $X_-^\dagger$ such that $X_+ X_-^\dagger$ is stable.

Proof 'if' part

Suppose that (a) and (b) hold.

Let $(A, B) \in \Sigma$.

$\implies X_+ = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \implies X_+ X_-^\dagger = A + BU_- X_-^\dagger$ is stable

$\implies K := U_- X_-^\dagger$ stabilizes (A, B).

Theorem stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback $\Longleftrightarrow$

(a) X_- has full row rank, and

(b) there exists **a** right inverse $X_-^\dagger$ such that $X_+X_-^\dagger$ is stable.

Proof 'only if' part

Suppose that $\exists\ K$ such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Let $(A, B) \in \Sigma$ and $\begin{bmatrix} A_0 & B_0 \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0$. Then, $\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = X_+$

Since $\begin{bmatrix} A + \alpha A_0 & B + \alpha B_0 \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = X_+, (A + \alpha A_0, B + \alpha B_0) \in \Sigma$ for all α.

Note that $\begin{bmatrix} (A_0 + B_0K)^\top A_0 & (A_0 + B_0K)^\top B_0 \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0.$

As such, $\Big(A + \alpha(A_0 + B_0K)^\top A_0, B + \alpha(A_0 + B_0K)^\top B_0\Big) \in \Sigma$ for all α.

Therefore, $M_\alpha := (A + BK) + \alpha(A_0 + B_0K)^\top(A_0 + B_0K)$ is stable for all α.

Let $\rho(M)$ denote the spectral radius of M, that is $\rho(M) := \max\{|\lambda| : \lambda \in \sigma(M)\}$.

Note that $0 = \lim_{\alpha\to\infty} \rho(\frac{M_\alpha}{\alpha}) = \rho\Big((A_0 + B_0K)^\top(A_0 + B_0K)\Big)$. Then, $A_0 + B_0K = 0$.

This proves that $\begin{bmatrix} A_0 & B_0 \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = 0 \implies \begin{bmatrix} A_0 & B_0 \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} = 0.$

As such, $\begin{bmatrix} I \\ K \end{bmatrix} = \begin{bmatrix} X_- \\ U_- \end{bmatrix} F$ for some F.

Then, F is a right inverse of X_- and $X_+F = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} F = A + BK$ is stable!

Theorem stabilization by state feedback

The data (U_-, X) are informative for stabilization by state feedback $\iff$

(a) X_- has full row rank, and

(b) there exists **a** right inverse $X_-^\dagger$ such that $X_+ X_-^\dagger$ is stable.

Theorem stabilization via linear matrix inequalities

The data (U_-, X) are informative for stabilization by state feedback $\iff$ there exists $\Theta \in \mathbb{R}^{T\times n}$ such that

$$X_-\Theta = (X_-\Theta)^\top \quad \text{and} \quad \begin{bmatrix} X_-\Theta & X_+\Theta \\ \Theta^\top X_+^\top & X_-\Theta \end{bmatrix} > 0. \tag{$\star$}$$

Moreover, $K = U_-\Theta(X_-\Theta)^{-1}$ is a stabilizing feedback for any Θ satisfying ($\star$).

Observation

Θ is an $\boldsymbol{T} \times n$ matrix. So, the LMIs ($\star$) have $n\boldsymbol{T}$ unknowns.

Theorem stabilization via linear matrix inequalities

The data (U_-, X) are informative for stabilization by state feedback $\iff$ there exists $\Theta \in \mathbb{R}^{T \times n}$ such that

$$X_-\Theta = (X_-\Theta)^\top \quad \text{and} \quad \begin{bmatrix} X_-\Theta & X_+\Theta \\ \Theta^\top X_+^\top & X_-\Theta \end{bmatrix} > 0. \qquad (\star)$$

Moreover, $K = U_-\Theta(X_-\Theta)^{-1}$ is a stabilizing feedback for any Θ satisfying $(\star)$.

Theorem stabilization via more friendly LMIs

The data (U_-, X) are informative for stabilization by state feedback $\iff$ there exists $P \in \mathbb{S}^n$, $L \in \mathbb{R}^{m \times n}$, and $\alpha > 0$ such that

$$P > 0 \quad \text{and} \quad \begin{bmatrix} P - \alpha I & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} + \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix} \begin{bmatrix} X_+ \\ -X_- \\ -U_- \\ 0 \end{bmatrix}^\top \geqslant 0. \qquad (\star\star)$$

Moreover, $K = LP^{-1}$ is a stabilizing feedback for any P, L satisfying $(\star\star)$.

Observation

The LMIs $(\star\star)$ have $\frac{1}{2}n(n+1) + nm + 1$ unknowns whereas those in $(\star)$ have $n\boldsymbol{T}$.

Recall optimal linear quadratic regulator

LQR(A, B, Q, R): Consider the discrete-time linear system

$$\boldsymbol{x}(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t), \qquad \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{u} \in \mathbb{R}^m \qquad (\star)$$

Notation: $\boldsymbol{x}_{x_0,\boldsymbol{u}}$ is the state sequence of $(\star)$ resulting from the input $\boldsymbol{u}$ and $\boldsymbol{x}(0) = x_0$. Whenever x_0 and $\boldsymbol{u}$ are clear from the context, we simply write $\boldsymbol{x}$.

Consider the quadratic cost functional

$$J(x_0, \boldsymbol{u}) = \sum_{t=0}^{\infty} \boldsymbol{x}^\top(t) Q \boldsymbol{x}(t) + \boldsymbol{u}^\top(t) R \boldsymbol{u}(t),$$

where $Q \in \mathbb{S}^n$ is positive semidefinite and $R \in \mathbb{S}^m$ is positive definite.

We say that LQR(A, B, Q, R) is solvable for (A, B, Q, R) if for every x_0 there exists an input $\boldsymbol{u}^*$ such that

(a) The cost $J(x_0, \boldsymbol{u}^*)$ is finite.

(b) The limit $\lim_{t\to\infty} \boldsymbol{x}_{x_0,\boldsymbol{u}^*}(t) = 0$.

(c) The input $\boldsymbol{u}^*$ minimizes the cost functional, i.e.,

$$J(x_0, \boldsymbol{u}^*) \leqslant J(x_0, \boldsymbol{u})$$

for all $\boldsymbol{u}$ such that $\lim_{t\to\infty} \boldsymbol{x}_{x_0,\boldsymbol{u}}(t) = 0$.

Theorem optimal linear quadratic regulator

$\mathrm{LQR}(A, B, Q, R)$ is solvable $\Longleftrightarrow$

(a) (A, B) is stabilizable and

(b) every eigenvalue of A on the unit circle is (Q, A)-observable, that is

$$\operatorname{rank}\begin{bmatrix} A - \lambda I \\ Q \end{bmatrix} = n \qquad \text{for all } \lambda \in \sigma(A) \text{ with } |\lambda| = 1.$$

If these conditions are satisfied, then

(i) there exists a unique largest positive semidefinite solution P^+ to the discrete-time algebraic Riccati equation

$$P = A^\top PA - A^\top PB(R + B^\top PB)^{-1}B^\top PA + Q, \qquad \text{(DARE)}$$

in the sense that $P^+ \geqslant P$ for every real symmetric P satisfying (DARE).

(ii) for every x_0 a unique optimal input $\boldsymbol{u}^*$ exists and is generated by the feedback law $\boldsymbol{u}^* = K\boldsymbol{x}$ where

$$K = -(R + B^\top P^+ B)^{-1} B^\top P^+ A. \qquad \text{(OFG)}$$

(iii) $A + BK$ is stable.

Definition

If $\mathrm{LQR}(A, B, Q, R)$ is solvable, we say that K given by (OFG) is the optimal feedback gain for (A, B, Q, R).

Recall data-consistent systems

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad \text{and} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

$$\Sigma = \{(A,B) : X_+ = AX_- + BU_-\} \qquad (A_s, B_s) \in \Sigma$$

Definition given $Q \geqslant 0$ and $R > 0$

The data (U_-, X) are informative for optimal linear quadratic regulation if there exists K such that K is the optimal feedback gain for LQR(A,B,Q,R) for all $(A,B) \in \Sigma$.

Theorem

The data (U_-, X) are informative for optimal linear quadratic regulation $\iff$ one of the following two conditions hold:

(a) (U_-, X) are informative for SI $(\Sigma = \{(A_s, B_s)\})$ and LQR(A_s, B_s, Q, R) is solvable.

(b) For all $(A,B) \in \Sigma$ we have $A = A_s$. Moreover, A_s is stable, $QA_s = 0$, and the optimal feedback gain is given by $K = 0$.

Theorem

The data (U_-, X) are informative for optimal linear quadratic regulation $\iff$ one of the following two conditions hold:

(a) (U_-, X) are informative for SI $(\Sigma = \{(A_s, B_s)\})$ and LQR(A_s, B_s, Q, R) is solvable.

(b) For all $(A, B) \in \Sigma$ we have $A = A_s$. Moreover, A_s is stable, $QA_s = 0$, and the optimal feedback gain is given by $K = 0$.

Remark

If (U_-, X) are informative for optimal linear quadratic regulation, there must exist $P^+_{(A,B)}$ satisfying the DARE

$$P^+_{(A,B)} = A^\top P^+_{(A,B)} A - A^\top P^+_{(A,B)} B (R + B^\top P^+_{(A,B)} B)^{-1} B^\top P^+_{(A,B)} A + Q$$

such that

$$K = -(R + B^\top P^+_{(A,B)} B)^{-1} B^\top P^+_{(A,B)} A.$$

for all $(A, B) \in \Sigma$.

It turns out that $P^+_{(A,B)}$ is independent of the choice of (A, B). A common solution for the DARE for all $(A, B) \in \Sigma$ can exist only if either (a) holds or the pathological case (b) holds.

Theorem

Define

$$\mathcal{L}(P) := X_-^\top P X_- - X_+^\top P X_+ - X_-^\top Q X_- - U_-^\top R U_-.$$

Then, (U_-, X) are informative for optimal linear quadratic regulation $\iff$ the optimization problem

$$\begin{aligned} &\text{maximize} && \operatorname{trace} P \\ &\text{subject to} && P = P^\top \geqslant 0 \text{ and } \mathcal{L}(P) \leqslant 0 \end{aligned}$$

has a unique solution. In that case, if P^+ is the unique solution, then there exists a right inverse of X_-, say $X_-^\dagger$, such that $\mathcal{L}(P^+)X_-^\dagger = 0$ and the optimal gain is given by $K = U_- X_-^\dagger$.

Exact data

Problem	D
stability	S
controllability	IS
stabilizability	IS
stability	IS
observability	SO
stabilization by state feedback	IS
deadbeat controller	IS
LQR	IS
suboptimal LQR	IS
suboptimal $\mathcal{H}_2$	IS
stabilization by dynamic feedback	ISO
stabilization by dynamic feedback	IO
dissipativity analysis	ISO
tracking and regulation	IS
model reduction (moment matching)	IO
reachability (conic constraints)	IO

I: input S: state O: output

Model-based vs. data-based systems and control theory

Data informativity framework

Data informativity with noise-free data (analysis)

Data informativity with noise-free data (design)

Data informativity with noisy data

Setup

noisy input-state data

model class $\boldsymbol{x}(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + \boldsymbol{w}(t)$ $\quad \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{u} \in \mathbb{R}^m, \boldsymbol{w} \in \mathbb{R}^n$

true system $\boldsymbol{x}(t+1) = A_s\boldsymbol{x}(t) + B_s\boldsymbol{u}(t) + \boldsymbol{w}(t)$ $\quad A_s, B_s, \boldsymbol{w}$ are unknown

noise model $W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$ is unknown but satisfies:

$$W_- W_-^\top \leqslant \varepsilon^2 I \tag{$\star$}$$

the data collected from the true system

$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

satisfy

$$X_+ = A_s X_- + B_s U_- + W_-$$

where

$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$$

and W_- is such that ($\star$) holds.

data-consistent systems $\quad (A_s, B_s) \in \Sigma$

$$\Sigma = \left\{ (A,B) : \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \varepsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

Data-based systems and control theory

Day 3

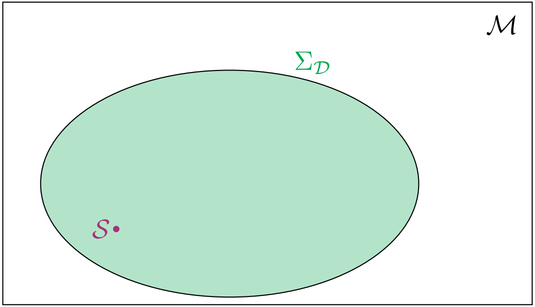
GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and **Henk van Waarde**

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen



$\mathcal{M}$: model class
$\mathcal{S}$: unknown system
$\mathcal{D}$: given data set
$\Sigma_{\mathcal{D}}$: data consistent systems

$\mathcal{P}$: system property
$\mathcal{O}$: control objective

- Data $\mathcal{D}$ are **informative for**
  - $\mathcal{P}$ $:\Longleftrightarrow$ all systems in $\Sigma_{\mathcal{D}}$ have property $\mathcal{P}$
  - $\mathcal{O}$ $:\Longleftrightarrow$ there exists a controller $\mathcal{C}$ that achieves $\mathcal{O}$ for all systems in $\Sigma_{\mathcal{D}}$.
- **Data driven control design** $:\Longleftrightarrow$ use $\mathcal{D}$ to find such $\mathcal{C}$

Consider the system

$$x(t+1) = A_s x(t) + B_s u(t) + w(t),$$

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input and $w(t) \in \mathbb{R}^n$ is the **process noise (not measured!)**

The matrices A_s and B_s are **unknown** but the following **data** are **given**:

$$\begin{aligned} X &:= \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \\ U_- &:= \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}. \end{aligned}$$

Goal: using the data (U_-, X), find a feedback law $u = Kx$ such that

$$x(t+1) = (A_s + B_s K)x(t)$$

is **asymptotically stable** (equivalently, $A_s + B_s K$ is a stable matrix).

Note: This is not possible without further assumptions on $w(t)$.

Assumption: The matrix

$$W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$$

is **unknown** but assumed to satisfy the **bound**

$$\|W_-\| \leqslant \epsilon,$$

where $\epsilon \geqslant 0$ is given and $\|\cdot\|$ denotes the spectral norm.

Recall: A symmetric matrix $P \in \mathbb{S}^n$ is **positive semidefinite** if $x^\top P x \geqslant 0$ for all $x \in \mathbb{R}^n$ and **positive definite** if $x^\top P x > 0$ for all nonzero $x \in \mathbb{R}^n$. We use the notation $P \geqslant 0$ and $P > 0$, respectively. Moreover, $P \geqslant Q$ means: $P - Q \geqslant 0$.

Fact: For any matrix Z, we have that $\|Z\| \leqslant \epsilon$ **if and only if**

$$ZZ^\top \leqslant \epsilon^2 I.$$

So our assumption on the noise is:

$$W_- W_-^\top \leqslant \epsilon^2 I.$$

Introduce the matrices:

$$X_- := \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix}, \quad X_+ := \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

Definition: The set Σ of **data-consistent systems** is given by:

$$\Sigma = \{(A,B) \mid X_+ = AX_- + BU_- + W_- \text{ for some } W_- \text{ satisfying } \|W_-\| \leqslant \epsilon\}.$$

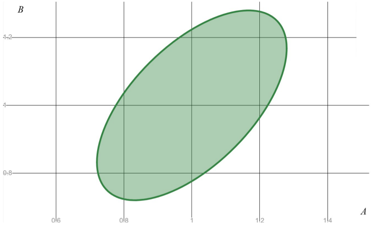
Example: Let $A_s = 1$, $B_s = 1$ and $\epsilon = \frac{1}{2}$.

Data: $X = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}$, $U_- = \begin{bmatrix} 1 & -2 \end{bmatrix}$.

Then Σ consists of all $A, B \in \mathbb{R}$ such that:

$$\left\| X_+ - \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \right\| \leqslant \epsilon,$$

i.e., $(2-(A+B))^2 + (2A-2B)^2 \leqslant 0.25$.



Set Σ

Definition: The data (U_-, X) are called **informative** for **quadratic stabilization** if there exists a feedback gain $K \in \mathbb{R}^{m\times n}$ and a $P \in \mathbb{S}^n$ such that $P > 0$ and

$$P - (A + BK)^\top P(A + BK) > 0 \qquad (*)$$

for all $(A, B) \in \Sigma$.

Interpretation: Informativity for quadratic stabilization means that $A + BK$ is **stable** for all $(A, B) \in \Sigma$ and

$$V(x) := x^\top Px$$

is a **common Lyapunov function** for all $(A, B) \in \Sigma$.

Indeed, $V(0) = 0$ and $V(x) > 0$ for all nonzero $x \in \mathbb{R}^n$. Moreover, due to $(*)$,

$$V(x(t+1)) = x(t)^\top (A + BK)^\top P(A + BK)x(t) < x(t)^\top Px(t) = V(x(t))$$

for all $x(t+1)$ and $x(t) \neq 0$ satisfying $x(t+1) = (A + BK)x(t)$.

Problem: Find conditions for informativity, and provide a K (if it exists).

Lemma: Let $S_{11} \in \mathbb{S}^q$, $S_{12} = S_{21}^\top \in \mathbb{R}^{q \times r}$ and $S_{22} \in \mathbb{S}^r$. Consider

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}.$$

1. $S > 0$ if and only if $S_{22} > 0$ and $S_{11} - S_{12} S_{22}^{-1} S_{21} > 0$.
2. $S > 0$ if and only if $S_{11} > 0$ and $S_{22} - S_{21} S_{11}^{-1} S_{12} > 0$.
3. Assume that $S_{22} > 0$. Then $S \geqslant 0$ if and only if $S_{11} - S_{12} S_{22}^{-1} S_{21} \geqslant 0$.
4. Assume that $S_{11} > 0$. Then $S \geqslant 0$ if and only if $S_{22} - S_{21} S_{11}^{-1} S_{12} \geqslant 0$.

Fact: Let $P \in \mathbb{S}^n$. We have that $P > 0$ if and only if $P^{-1} > 0$.

Thus, by 1, $P > 0$ and $P - (A+BK)^\top P (A+BK) > 0$ if and only if

$$\begin{bmatrix} P & (A+BK)^\top \\ A+BK & P^{-1} \end{bmatrix} > 0.$$

Equivalently, by 2, $P^{-1} > 0$ and $P^{-1} - (A+BK) P^{-1} (A+BK)^\top > 0$.

Conclusion: The data (U_-, X) are informative for quadratic stabilization **if and only if** there exists $K \in \mathbb{R}^{m\times n}$ and $P \in \mathbb{S}^n$ such that $P > 0$ and

$$P - (A + BK)P(A + BK)^\top > 0 \qquad (**)$$

for all $(A, B) \in \Sigma$.

Note: Since $A + BK = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix}$ we can rewrite $(**)$ as

$$\left[\begin{array}{c|cc} I & A & B \end{array}\right] \left[\begin{array}{c|c} P & 0 \\ \hline 0 & -\begin{bmatrix} I \\ K \end{bmatrix} P \begin{bmatrix} I \\ K \end{bmatrix}^\top \end{array}\right] \left[\begin{array}{c} I \\ \hline A^\top \\ B^\top \end{array}\right] > 0.$$

equivalently,

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0.$$

This is a **quadratic matrix inequality (QMI)** in A and B.

Recall: The set Σ of **data-consistent systems** equals:

$$\Sigma = \left\{ (A,B) \mid \left\| X_+ - \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} \right\| \leqslant \epsilon \right\}.$$

The norm constraint is equivalent to

$$\begin{bmatrix} I & A & B \end{bmatrix} \begin{bmatrix} X_+ \\ -X_- \\ -U_- \end{bmatrix} \begin{bmatrix} X_+ \\ -X_- \\ -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \leqslant \epsilon^2 I.$$

Since $\epsilon^2 I = \begin{bmatrix} I & A & B \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}$, we can reformulate this as

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0.$$

This is yet another **quadratic matrix inequality** in A and B!

The **main question** is thus:
When do there exist $P > 0$ and K such that **all systems** (A, B) satisfying

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0$$

also satisfy the QMI

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0 \, ?$$

Assume that there exist matrices $K \in \mathbb{R}^{m\times n}$ and $P \in \mathbb{S}^n$, and a scalar $\alpha \geqslant 0$ such that $P > 0$ and

$$\underbrace{\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix}}_{=:M(P,K)} - \alpha \underbrace{\begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top}_{=:N} > 0.$$

Let $(A, B) \in \Sigma$. Then $\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top N \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0.$

Therefore,

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top M(P,K) \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > \alpha \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top N \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0.$$

Thus, the data (U_-, X) are **informative for quadratic stabilization**!

Proposition: The data (U_-, X) are **informative for quadratic stabilization** if there exist $K \in \mathbb{R}^{m\times n}$, $P \in \mathbb{S}^n$ and a scalar $\alpha \geqslant 0$ such that $P > 0$ and

$$\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0. \quad (\triangle)$$

Moreover, K is such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Example: Let $\epsilon = \frac{1}{2}$ and consider the **data** $X = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}$, $U_- = \begin{bmatrix} 1 & -2 \end{bmatrix}$. One can verify that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} - 1 \begin{bmatrix} -3.75 & 2 & 2 \\ 2 & -5 & 3 \\ 2 & 3 & -5 \end{bmatrix} > 0.$$

Therefore, $(\triangle)$ holds with $K = -1$, $P = 1$ and $\alpha = 1$.

Hence the data (U_-, X) are **informative for quadratic stabilization**. The controller $u = -x$ stabilizes all systems $(A, B) \in \Sigma$.

The matrix inequality is **nonlinear** in P and K. How to compute a solution?

Proposition: The data (U_-, X) are **informative for quadratic stabilization** if there exist matrices $L \in \mathbb{R}^{m\times n}$ and $P \in \mathbb{S}^n$, and a scalar $\alpha \geqslant 0$ such that

$$\begin{bmatrix} P & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix}^\top > 0.$$

The feedback gain $K = LP^{-1}$ stabilizes all systems in Σ.

Note: The condition is now a **linear matrix inequality**!

Proof: A Schur complement w.r.t. the $(4,4)$-block yields $P > 0$ and

$$\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -L^\top \\ 0 & -L & -LP^{-1}L^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0.$$

Now, define $K := LP^{-1}$ and conclude that the data are informative. $\square$

Proposition: The data (U_-, X) are informative for quadratic stabilization **if there exist** $K \in \mathbb{R}^{m\times n}$, $P \in \mathbb{S}^n$ and a scalar $\alpha \geqslant 0$ such that $P > 0$ and

$$\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0.$$

Moreover, K is such that $A + BK$ is stable for all $(A, B) \in \Sigma$.

Question: Are these conditions also **necessary** for informativity?

Theorem[1]: Let $M, N \in \mathbb{S}^n$. Then $x^\top M x > 0$ for all nonzero $x \in \mathbb{R}^n$ satisfying $x^\top N x \geqslant 0$ **if and only if** there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

V.A. Yakubovich
Russian control theorist
(1926 – 2012)

Remarks:

- Originally developed in the context of absolute stability of **Lur'e systems**.
- Non-strict version available.
- But for our purposes we need a **matrix version of the S-lemma!**

[1] V. A. Yakubovich, "S-procedure in nonlinear control theory," *Vestnik Leningrad University Mathematics*, 1977.

SIAM J. CONTROL OPTIM.
Vol. 61, No. 4, pp. 2251–2281

QUADRATIC MATRIX INEQUALITIES WITH APPLICATIONS TO DATA-BASED CONTROL*

HENK J. VAN WAARDE[†], M. KANAT CAMLIBEL[†], JAAP EISING[‡], AND HARRY L. TRENTELMAN[‡]

Abstract. This paper studies several problems related to quadratic matrix inequalities (QMIs), i.e., inequalities in the Loewner order involving quadratic functions of matrix variables. In particular, we provide conditions under which the solution set of a QMI is nonempty, convex, or bounded or has a nonempty interior. We also provide a parameterization of the solution set of a given QMI.

Figure: The “cookbook" for data driven control problems with noise.

Let $M, N \in \mathbb{S}^{q+r}$ be partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \quad N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}.$$

Theorem: Assume that $N_{22} < 0$ and $N_{11} - N_{12}N_{22}^{-1}N_{21} \geqslant 0$. Then

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \text{ for all } Z \in \mathbb{R}^{r\times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

Remark: Note that

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} = N_{11} - N_{12}N_{22}^{-1}N_{21} + (N_{22}^{-1}N_{21} + Z)^\top N_{22}(N_{22}^{-1}N_{21} + Z).$$

Thus, assuming $N_{22} < 0$, the condition $N_{11} - N_{12}N_{22}^{-1}N_{21} \geqslant 0$ is equivalent to:

$$\exists Z \in \mathbb{R}^{r\times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0.$$

Recall: The data (U_-, X) are **informative for quadratic stabilization** if and only if there exist $P > 0$ and K such that

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0$$

holds **for all systems** (A, B) satisfying

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0.$$

Theorem: Assume that

$$\operatorname{rank}\begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m. \qquad \text{(RankCond)}$$

Then the data (U_-, X) are informative for quadratic stabilization **if and only if** there exist $K \in \mathbb{R}^{m \times n}$, $P \in \mathbb{S}^n$ and a scalar $\alpha \geqslant 0$ such that $P > 0$ and

$$\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0.$$

The feedback gain K stabilizes all systems in Σ.

Proof: Define

$$N = \left[\begin{array}{c|c} N_{11} & N_{12} \\ \hline N_{21} & N_{22} \end{array}\right] = \left[\begin{array}{cc} I & X_+ \\ \hline 0 & -X_- \\ 0 & -U_- \end{array}\right] \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \left[\begin{array}{cc} I & X_+ \\ \hline 0 & -X_- \\ 0 & -U_- \end{array}\right]^\top .$$

By (RankCond), $N_{22} = -\begin{bmatrix} X_- \\ U_- \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix}^\top < 0$. Also, $\begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$ for $Z = \begin{bmatrix} A_s^\top \\ B_s^\top \end{bmatrix}$. The theorem now follows from the **matrix S-lemma**. $\square$

Stabilization using noisy data

A sufficient condition

Classical S-lemma

The matrix S-lemma

Necessary and sufficient conditions

Generalization of the noise model

Generalization to rank-deficient data

Recall: The assumption on the noise was: $W_- W_-^\top \leqslant \epsilon^2 I$, equivalently,

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \begin{bmatrix} \epsilon^2 I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0.$$

We can also work with the (more general) bound:

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}}_{=:\Phi} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0,$$

for a known matrix $\Phi \in \mathbb{S}^{n+T}$.

Special cases:

- Energy bound: If $\Phi_{12} = 0$ and $\Phi_{22} = -I$ then $\sum_{t=0}^{T-1} w(t)w(t)^\top \leqslant \Phi_{11}$.
- Norm bound: If $\Phi_{11} = \epsilon^2 I$, $\Phi_{12} = 0$ and $\Phi_{22} = -I$ we recover $W_- W_-^\top \leqslant \epsilon^2 I$.
- Noiseless case: If, in addition, $\epsilon = 0$ then $W_- = 0$.
- Sample bounds: If $\|w(t)\| \leqslant \epsilon \; \forall t$ then energy bound holds with $\Phi_{11} = \epsilon^2 T I$.
- Sample covariance bounds: If $\Phi_{22} = \frac{1}{T}(-I + \frac{1}{T}\mathbb{1}\mathbb{1}^\top)$, $\Phi_{12} = 0$ and $\mu = \frac{1}{T}\sum_{t=0}^{T-1} w(t)$ then $\frac{1}{T}\sum_{t=0}^{T-1}(w(t) - \mu)(w(t) - \mu)^\top \leqslant \Phi_{11}$.

Question: What changes to our conditions for informativity?

Answer: Luckily, not much!

Theorem: Suppose that the noise matrix W_- satisfies

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0.$$

Assume that $\Phi_{22} < 0$ and

$$\operatorname{rank} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m.$$

Then the data (U_-, X) are informative for quadratic stabilization **if and only if** there exist $K \in \mathbb{R}^{m \times n}$, $P \in \mathbb{S}^n$ and a scalar $\alpha \geqslant 0$ such that $P > 0$ and

$$\begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0.$$

We consider one final theorem from the “cookbook".

Theorem: Assume that $N_{22} \leqslant 0$, $\ker N_{22} \subseteq \ker N_{21}$ and $N_{11} - N_{12}N_{22}^{\dagger}N_{21} \geqslant 0$. Moreover, assume that $M_{22} \leqslant 0$. Then

$$\begin{bmatrix} I \\ Z \end{bmatrix}^{\top} M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \text{ for all } Z \in \mathbb{R}^{r \times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^{\top} N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exist scalars $\alpha \geqslant 0$ and $\beta > 0$ such that

$$M - \alpha N \geqslant \begin{bmatrix} \beta I & 0 \\ 0 & 0 \end{bmatrix}.$$

Theorem: Assume that $\Phi_{22} \leqslant 0$, $\ker \Phi_{22} \subseteq \ker \Phi_{21}$ and $\Phi_{11} - \Phi_{12}\Phi_{22}^{\dagger}\Phi_{21} \geqslant 0$. Then the data (U_-, X) are informative for quadratic stabilization **if and only if** there exist $K \in \mathbb{R}^{m \times n}$, $P \in \mathbb{S}^n$ and scalars $\alpha \geqslant 0$ and $\beta > 0$ such that $P > 0$ and

$$\begin{bmatrix} P - \beta I & 0 & 0 \\ 0 & -P & -PK^{\top} \\ 0 & -KP & -KPK^{\top} \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^{\top} \geqslant 0.$$

- **Setting**: model class of noisy input-state systems with noise model Φ.
- **Data**: input-state samples (U_-, X).
- **Check feasibility** of the **linear matrix inequality**

$$\begin{bmatrix} P-\beta I & 0 & 0 & 0 \\ 0 & -P & -L^\top & 0 \\ 0 & -L & 0 & L \\ 0 & 0 & L^\top & P \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \end{bmatrix}^\top \geqslant 0.$$

  Solution: $P \in \mathbb{S}^n$, $L \in \mathbb{R}^{m \times n}$ and scalars $\alpha \geqslant 0$ and $\beta > 0$.
- The **feedback gain** $K = LP^{-1}$ stabilizes all $(A, B) \in \Sigma$.

Compared to related results in [2,3]:

- **non-conservative** control design
  - if and only if conditions for informativity via matrix S-lemma
- **low complexity**
  - Dimensions of the unknown matrices are **independent of T**
  - Size of the matrix inequality is $(3n+m) \times (3n+m)$, also independent of T.

[2] J. Berberich, A. Koch, C. W. Scherer, F Allgöwer, "Robust data driven state-feedback design", *ACC*, 2020.

[3] C. De Persis and P. Tesi, "Formulas for Data-driven Control: Stabilization, Optimality and Robustness", *IEEE TAC*, 2020.

Data-based systems and control theory

Day 4

GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and Henk van Waarde

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen

SIAM J. CONTROL OPTIM.
Vol. 61, No. 4, pp. 2251–2281

QUADRATIC MATRIX INEQUALITIES WITH APPLICATIONS TO DATA-BASED CONTROL*

HENK J. VAN WAARDE†, M. KANAT CAMLIBEL†, JAAP EISING‡, AND HARRY L. TRENTELMAN‡

Abstract. This paper studies several problems related to quadratic matrix inequalities (QMIs), i.e., inequalities in the Loewner order involving quadratic functions of matrix variables. In particular, we provide conditions under which the solution set of a QMI is nonempty, convex, or bounded or has a nonempty interior. We also provide a parameterization of the solution set of a given QMI.

Figure: The "cookbook" for data driven control problems with noise.

Today's menu:

- beyond stabilization — $\mathcal{H}_\infty$ control
- dualization — dissipativity analysis
- projection — stabilization from IO data

Recap

$\mathcal{H}_\infty$ control

Dissipativity analysis

Stabilization from input-output data

Summary and outlook

Setup

noisy input-state data

model class $x(t+1) = Ax(t) + Bu(t) + w(t)$ $\quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, w \in \mathbb{R}^n$

true system $x(t+1) = A_s x(t) + B_s u(t) + w(t)$ $\quad A_s, B_s, w$ are unknown

noise model $W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$ is unknown but satisfies:

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0$$

where

$$\Phi_{11} \in \mathbb{S}^n, \quad \Phi_{22} \in \mathbb{S}^T, \quad \text{and} \quad \Phi_{21} = \Phi_{12}^\top.$$

Recall

data-consistent systems

$$\Sigma = \left\{ (A,B) : \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

Recall data-consistent systems

$$\Sigma = \left\{ (A,B) : \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

Definition stabilization by state feedback

The data (U_-, X) are informative for quadratic stabilization if there exist a feedback gain K and a matrix $P = P^\top > 0$ such that for all $(A,B) \in \Sigma$

$$P - (A+BK)P(A+BK)^\top > 0. \qquad (\star)$$

Observation

$(\star)$ is equivalent to
$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0$$

Question

one QMI implies another

Under what conditions is the QMI

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0$$

satisfied by for all (A, B) such that

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \quad ?$$

Theorem

Let $M, N \in \mathbb{S}^{q+r}$. Partition N as $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$. Assume that $N_{22} < 0$ and $N_{11} - N_{12}N_{22}^{-1}N_{21} \geqslant 0$. Then

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \text{ for all } Z \in \mathbb{R}^{r\times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

Recap

$\mathcal{H}_\infty$ control

Dissipativity analysis

Stabilization from input-output data

Summary and outlook

Terminology

from square integrable sequences to $\mathcal{H}_\infty$ performance

- $\ell_2(\mathbb{Z}_+, \mathbb{R}^q)$: the space of all $\mathbb{R}^q$-valued sequences v such that $\sum_{t=0}^{\infty} \|v(t)\|^2 < \infty$.
- For $v \in \ell_2(\mathbb{Z}_+, \mathbb{R}^q)$, we define its ℓ_2-norm by $\|v\|_{\ell_2} := \left(\sum_{t=0}^{\infty} \|v(t)\|^2\right)^{\frac{1}{2}}$.
- Consider the input-state-output system

$$
\begin{aligned}
x(t+1) &= Ax(t) + w(t) \\
z(t) &= Cx(t)
\end{aligned}
$$

  where $w(t) \in \mathbb{R}^q$ and $z(t) \in \mathbb{R}^p$.
- If A is stable and $x(0) = 0$, then $z \in \ell_2(\mathbb{Z}_+, \mathbb{R}^p)$ whenever $w \in \ell_2(\mathbb{Z}_+, \mathbb{R}^q)$ and the $\mathcal{H}_\infty$ performance of the system is defined as

$$
J_{\mathcal{H}_\infty} = \sup_{\|w\|_{\ell_2}=1} \|z\|_{\ell_2}.
$$

- Discrete-time bounded real lemma: Let $\gamma > 0$. Then, A is stable and $J_{\mathcal{H}_\infty} < \gamma$ if and only if there exists $P > 0$ such that

$$
\begin{bmatrix} P - A^\top P A - C^\top C & -A^\top P \\ -PA & \gamma^2 I - P \end{bmatrix} > 0.
$$

Problem state feedback $\mathcal{H}_\infty$ control

Given the system

$$x(t+1) = Ax(t) + Bu(t) + w(t)$$
$$z(t) = Cx(t) + Du(t),$$

find a state feedback $u = Kx$ such that $A + BK$ is stable and the $\mathcal{H}_\infty$ performance of closed-loop system

$$x(t+1) = (A + BK)x(t) + w(t)$$
$$z(t) = (C + DK)x(t)$$

is less than $\gamma > 0$.

Theorem

The state feedback $\mathcal{H}_\infty$ control is solvable if and only if there exist $P > 0$ and K such that

$$\begin{bmatrix} P - (A+BK)^\top P(A+BK) - (C+DK)^\top (C+DK) & -(A+BK)^\top P \\ -P(A+BK) & \gamma^2 I - P \end{bmatrix} > 0.$$

Theorem

The state feedback $\mathcal{H}_\infty$ control is solvable if and only if there exist $P>0$ and K such that

$$\begin{bmatrix} P-(A+BK)^\top P(A+BK)-(C+DK)^\top(C+DK) & -(A+BK)^\top P \\ -P(A+BK) & \gamma^2 I - P \end{bmatrix} > 0. \qquad (\star)$$

Observation

The matrix inequality $(\star)$ is **not** a QMI in (A,B):

$$(A+BK)^\top P(A+BK) = A^\top PA + A^\top PBK + K^\top B^\top PA + K^\top B^\top PBK.$$

Idea $A_K = A+BK$ and $C_K = C+DK$

Employ standard matrix acrobatics from LMI theory:

$$\begin{bmatrix} P - A_K^\top P A_K - C_K^\top C_K & -A_K^\top P \\ -PA_K & \gamma^2 I - P \end{bmatrix} \overset{(\star)}{>} 0 \overset{\text{Schur comp.}}{\iff}$$

$$\gamma^2 I - P > 0 \qquad \text{and} \qquad P - A_K^\top\left(P + P(\gamma^2 I - P)^{-1}P\right)A_K - C_K^\top C_K > 0$$

Idea $A_K = A + BK$, $C_K = C + DK$, and $Q = P^{-1}$

$$\gamma^2 I - P > 0 \quad \overset{P^{-\frac{1}{2}}\,\bullet}{\underset{\bullet\,P^{-\frac{1}{2}}}{\iff}} \quad Q - \gamma^{-2} I > 0$$

$$P - A_K^\top \left(P + P(\gamma^2 I - P)^{-1} P\right) A_K - C_K^\top C_K > 0$$

$$\Updownarrow$$

$$P - A_K^\top (Q - \gamma^{-2} I)^{-1} A_K - C_K^\top C_K > 0 \quad \overset{\text{Schur comp.}}{\iff} \quad \begin{bmatrix} P - C_K^\top C_K & A_K^\top \\ A_K & Q - \gamma^{-2} I \end{bmatrix} > 0$$

$$\overset{\left[\begin{smallmatrix} Q & \\ & I \end{smallmatrix}\right]\bullet}{\underset{\bullet\left[\begin{smallmatrix} Q & \\ & I \end{smallmatrix}\right]}{\iff}} \quad \begin{bmatrix} Q - (CQ + DKQ)^\top (CQ + DKQ) & Q(A^\top + K^\top B^\top) \\ AQ + BKQ & Q - \gamma^{-2} I \end{bmatrix} > 0$$

$$\overset{L = KQ}{\iff} \quad \begin{bmatrix} Q - (CQ + DL)^\top (CQ + DL) & QA^\top + L^\top B^\top \\ AQ + BL & Q - \gamma^{-2} I \end{bmatrix} > 0$$

Idea $C_{Q,L} = CQ + DL$

$$\begin{bmatrix} Q - (CQ+DL)^\top (CQ+DL) & QA^\top + L^\top B^\top \\ AQ + BL & Q - \gamma^{-2} I \end{bmatrix} > 0$$

$$\overset{\text{Schur comp.}}{\Longleftrightarrow}$$

$$Q - C_{Q,L}^\top C_{Q,L} > 0$$

and

$$Q - \gamma^{-2} I - (AQ + BL)(Q - C_{Q,L}^\top C_{Q,L})^{-1}(QA^\top + L^\top B^\top) > 0$$

Observation

The last inequality is a QMI in (A, B):

$$\begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix}^\top \left[\begin{array}{c:c} Q - \gamma^{-2} I & 0 \\ \hdashline 0 & -\begin{bmatrix} Q \\ L \end{bmatrix} (Q - C_{Q,L}^\top C_{Q,L})^{-1} \begin{bmatrix} Q \\ L \end{bmatrix}^\top \end{array}\right] \begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix} > 0$$

Setup

noisy input-state data

true system
$$x(t+1) = A_s x(t) + B_s u(t) + w(t)$$
A_s, B_s, w are unknown
$$z(t) = Cx(t) + Du(t)$$
C, D are given

noise model $W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \end{bmatrix}$ is unknown but satisfies:
$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0 \qquad (\star)$$

the data collected from the true system
$$X = \begin{bmatrix} x(0) & x(1) & \cdots & x(T) \end{bmatrix} \quad \text{and} \quad U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$
satisfy
$$X_+ = A_s X_- + B_s U_- + W_-$$
where
$$X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix} \quad X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}.$$

data-consistent systems $(A_s, B_s) \in \Sigma$
$$\Sigma = \left\{ (A,B) : \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \right\}.$$

Definition stabilization by state feedback

The data (U_-, X) are informative for $\mathcal{H}_\infty$ control with performance γ if there exist a feedback gain K and a matrix $P > 0$ such that for all $(A, B) \in \Sigma$

$$\begin{bmatrix} P - (A+BK)^\top P(A+BK) - (C+DK)^\top (C+DK) & -(A+BK)^\top P \\ -P(A+BK) & \gamma^2 I - P \end{bmatrix} > 0. \quad (1)$$

Observation

$\exists\ P > 0$ and K s.t. (1) holds if and only if $\exists\ Q > 0$ and L s.t. $Q - \gamma^{-2} I > 0$ and

$$\begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix}^\top \left[\begin{array}{c:c} Q - \gamma^{-2} I & 0 \\ \hdashline 0 & -\begin{bmatrix} Q \\ L \end{bmatrix} (Q - C_{Q,L}^\top C_{Q,L})^{-1} \begin{bmatrix} Q \\ L \end{bmatrix}^\top \end{array}\right] \begin{bmatrix} I \\ \hdashline A^\top \\ B^\top \end{bmatrix} > 0.$$

Recall

The QMI

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 \\ 0 & -P & -PK^\top \\ 0 & -KP & -KPK^\top \end{bmatrix} \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} > 0$$

needs to be satisfied for all $(A, B) \in \Sigma$ such that

$$\begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix}^\top \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top \begin{bmatrix} I \\ A^\top \\ B^\top \end{bmatrix} \geqslant 0 \quad ?$$

Theorem

Let $M, N \in \mathbb{S}^{q+r}$. Partition N as $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$. Assume that $N_{22} < 0$ and $N_{11} - N_{12}N_{22}^{-1}N_{21} \geqslant 0$. Then

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \text{ for all } Z \in \mathbb{R}^{r \times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

Theorem

Suppose that $\operatorname{rank} \begin{bmatrix} X_- \\ U_- \end{bmatrix} = n + m$ and $\Phi_{22} < 0$. Then, the data (U_-, X) are informative for $\mathcal{H}_\infty$ control with performance γ if there exist $Q > 0$, L, and $\alpha > 0$ s.t. $Q - \gamma^{-2} I > 0$ and

$$\begin{bmatrix} Q - \gamma^{-2} I & 0 \\ 0 & - \begin{bmatrix} Q \\ L \end{bmatrix} (Q - C_{Q,L}^\top C_{Q,L})^{-1} \begin{bmatrix} Q \\ L \end{bmatrix}^\top \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \end{bmatrix}^\top > 0.$$

Observation LMI formulation

$$\begin{bmatrix} Q - \gamma^{-2} I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Q & 0 \\ 0 & 0 & 0 & L & 0 \\ 0 & Q & L^\top & Q & C_{Q,L}^\top \\ 0 & 0 & 0 & C_{Q,L} & I \end{bmatrix} - \alpha \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & X_+ \\ 0 & -X_- \\ 0 & -U_- \\ 0 & 0 \\ 0 & 0 \end{bmatrix}^\top > 0.$$

Definition dissipativity

Consider a discrete-time linear input-state-output system:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \qquad (\star)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $y \in \mathbb{R}^p$. Let $S \in \mathbb{S}^{m+p}$.

The system $(\star)$ is said to be dissipative with respect to the supply rate

$$s(u, y) = \begin{bmatrix} u \\ y \end{bmatrix}^\top S \begin{bmatrix} u \\ y \end{bmatrix}$$

if there exists $P \in \mathbb{S}^n$ with $P \geqslant 0$ such that the dissipation inequality

$$x(t)^\top P x(t) + s\big(u(t), y(t)\big) \geqslant x(t+1)^\top P x(t+1)$$

holds for all $t \geqslant 0$ and for all trajectories $(u, x, y) : \mathbb{Z}_+ \to \mathbb{R}^{m+n+p}$ of $(\star)$.

If P satisfies the dissipation inequality that the function $x \mapsto x^\top P x$ is called a storage function.

Example positive-real and bounded-real

Positive-real case ($m = p$ and $S = \left[\begin{smallmatrix} 0 & I_m \\ I_m & 0 \end{smallmatrix}\right]$) and bounded-real case ($S = \left[\begin{smallmatrix} \gamma^2 I_m & 0 \\ 0 & -I_p \end{smallmatrix}\right]$)

Discussion — from dissipation inequality to an LMI

Along all trajectories of the system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \qquad (\star)$$

the dissipation inequality

$$x(t)^\top P x(t) + \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}^\top S \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} \geqslant x(t+1)^\top P x(t+1)$$

must hold for some $P \geqslant 0$. Rewrite the dissipation inequality:

$$\begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^\top \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} + \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^\top \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top S \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \geqslant \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^\top \begin{bmatrix} I & 0 \\ A & B \end{bmatrix}^\top \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$$

This leads to:

$$\begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^\top \left(\begin{bmatrix} I & 0 \\ A & B \end{bmatrix}^\top \begin{bmatrix} P & 0 \\ 0 & -P \end{bmatrix} \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} + \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top S \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \right) \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \geqslant 0.$$

Theorem — an LMI characterization

The system $(\star)$ is dissipative w.r.t. the supply rate s if and only if there exists $P \geqslant 0$ s.t.

$$\begin{bmatrix} I & 0 \\ A & B \end{bmatrix}^\top \begin{bmatrix} P & 0 \\ 0 & -P \end{bmatrix} \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} + \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top S \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \geqslant 0.$$

Setup

noisy input-state-output data

model class
$$x(t+1) = Ax(t) + Bu(t) + w(t) \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^m, w \in \mathbb{R}^n$$
$$y(t) = Cx(t) + Du(t) + z(t) \qquad y \in \mathbb{R}^p, z \in \mathbb{R}^p$$

true system
$$x(t+1) = A_s x(t) + B_s u(t) + w(t) \qquad A_s,\ B_s,\ w \text{ are unknown}$$
$$y(t) = C_s x(t) + D_s u(t) + z(t) \qquad C_s,\ D_s,\ z \text{ are unknown}$$

data $U_- = \begin{bmatrix} u(0) & \cdots & u(T-1) \end{bmatrix}$ $X = \begin{bmatrix} x(0) & \cdots & x(T) \end{bmatrix}$ $Y_- = \begin{bmatrix} y(0) & \cdots & y(T-1) \end{bmatrix}$

notation $X_- = \begin{bmatrix} x(0) & x(1) & \cdots & x(T-1) \end{bmatrix}$ $X_+ = \begin{bmatrix} x(1) & x(2) & \cdots & x(T) \end{bmatrix}$

noise model $V_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-1) \\ z(0) & z(1) & \cdots & z(T-1) \end{bmatrix}$ is unknown but satisfies:

$$\begin{bmatrix} I \\ V_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ V_-^\top \end{bmatrix} \geqslant 0 \qquad (\Phi_{11} \in \mathbb{S}^{n+p} \quad \text{and} \quad \Phi_{22} \in \mathbb{S}^{T})$$

data-consistent systems

$$\Sigma = \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} X_+ \\ Y_- \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X_- \\ U_- \end{bmatrix} + V_- \text{ s.t. } \begin{bmatrix} I \\ V_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ V_-^\top \end{bmatrix} \geqslant 0 \right\}$$

Setup

noisy input-state-output data

data-consistent systems

$$\Sigma = \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix}^\top \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right] \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right]^\top \begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix} \geqslant 0 \right\}$$

Definition

dissipativity

(U_-, X, Y_-) are informative for dissipativity w.r.t. the supply rate s if there exists $P \geqslant 0$ s.t.

$$\begin{bmatrix} I & 0 \\ A & B \end{bmatrix}^\top \begin{bmatrix} P & 0 \\ 0 & -P \end{bmatrix} \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} + \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^\top S \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \geqslant 0 \qquad (\star)$$

for all $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \Sigma$.

Observation

Let $S = \begin{bmatrix} S_{uu} & S_{uy} \\ S_{yu} & S_{yy} \end{bmatrix}$. Then, $(\star) \iff \begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 & 0 \\ 0 & S_{uu} & 0 & S_{uy} \\ 0 & 0 & -P & 0 \\ 0 & S_{yu} & 0 & S_{yy} \end{bmatrix} \begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix} \geqslant 0.$

Observation

The QMI

$$\begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 & 0 \\ 0 & S_{uu} & 0 & S_{uy} \\ 0 & 0 & -P & 0 \\ 0 & S_{yu} & 0 & S_{yy} \end{bmatrix} \begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix} \geqslant 0$$

needs to be satisfied for all $(A, B) \in \Sigma$ such that

$$\begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix}^\top \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right] \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right]^\top \begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix} \geqslant 0.$$

Theorem

strict matrix S-lemma

Let $M, N \in \mathbb{S}^{q+r}$. Partition N as $N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$. Assume that $N_{22} < 0$ and $N_{11} - N_{12}N_{22}^{-1}N_{21} \geqslant 0$. Then

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} > 0 \text{ for all } Z \in \mathbb{R}^{r\times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N > 0$.

SIAM J. CONTROL OPTIM.
Vol. 61, No. 4, pp. 2251–2281

QUADRATIC MATRIX INEQUALITIES WITH APPLICATIONS TO DATA-BASED CONTROL*

HENK J. VAN WAARDE†, M. KANAT CAMLIBEL†, JAAP EISING‡, AND HARRY L. TRENTELMAN‡

Abstract. This paper studies several problems related to quadratic matrix inequalities (QMIs), i.e., inequalities in the Loewner order involving quadratic functions of matrix variables. In particular, we provide conditions under which the solution set of a QMI is nonempty, convex, or bounded or has a nonempty interior. We also provide a parameterization of the solution set of a given QMI.

Figure: The “cookbook” for data driven control problems with noise.

Theorem dualization

Let $\Pi \in \mathbb{S}^{q+r}$. Suppose that Π is nonsingular. Then,

$$\begin{bmatrix} I_q \\ Z \end{bmatrix}^\top \Pi \begin{bmatrix} I_q \\ Z \end{bmatrix} \geqslant 0 \iff \begin{bmatrix} I_r \\ Z^\top \end{bmatrix}^\top \Pi^\sharp_{r,q} \begin{bmatrix} I_r \\ Z^\top \end{bmatrix} \geqslant 0$$

where

$$\Pi^\sharp_{r,q} := \begin{bmatrix} 0 & -I_r \\ I_q & 0 \end{bmatrix} \Pi^{-1} \begin{bmatrix} 0 & -I_q \\ I_r & 0 \end{bmatrix}.$$

Observation

Suppose that both P and S are nonsingular, then

$$\begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix}^\top \begin{bmatrix} P & 0 & 0 & 0 \\ 0 & S_{uu} & 0 & S_{uy} \\ 0 & 0 & -P & 0 \\ 0 & S_{yu} & 0 & S_{yy} \end{bmatrix} \begin{bmatrix} I \\ \hline A & B \\ C & D \end{bmatrix} \geqslant 0$$

$$\Updownarrow$$

$$\begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix}^\top \begin{bmatrix} P^{-1} & 0 & 0 & 0 \\ 0 & R_{yy} & 0 & R_{yu} \\ 0 & 0 & -P^{-1} & 0 \\ 0 & R_{uy} & 0 & R_{uu} \end{bmatrix} \begin{bmatrix} I \\ \hline A^\top & C^\top \\ B^\top & D^\top \end{bmatrix} \geqslant 0$$

where

$$\begin{bmatrix} R_{yy} & R_{yu} \\ R_{uy} & R_{uu} \end{bmatrix} = \begin{bmatrix} 0 & -I_p \\ I_m & 0 \end{bmatrix} \begin{bmatrix} S_{uu} & S_{uy} \\ S_{yu} & S_{yy} \end{bmatrix}^{-1} \begin{bmatrix} 0 & -I_m \\ I_p & 0 \end{bmatrix}.$$

SIAM J. CONTROL OPTIM. Vol. 61, No. 4, pp. 2251–2281 © 2023 Society for Industrial and Applied Mathematics

QUADRATIC MATRIX INEQUALITIES WITH APPLICATIONS TO DATA-BASED CONTROL*

HENK J. VAN WAARDE†, M. KANAT CAMLIBEL†, JAAP EISING‡, AND HARRY L. TRENTELMAN‡

Abstract. This paper studies several problems related to quadratic matrix inequalities (QMIs), i.e., inequalities in the Loewner order involving quadratic functions of matrix variables. In particular, we provide conditions under which the solution set of a QMI is nonempty, convex, or bounded or has a nonempty interior. We also provide a parameterization of the solution set of a given QMI.

Figure: The "cookbook" for data driven control problems with noise.

Theorem **nonstrict** matrix S-lemma

Let $M, N \in \mathbb{S}^{q+r}$. Assume that

- $N_{22} \leqslant 0$,
- $N_{11} - N_{12}N_{22}^{-1}N_{21} \gneq 0$, and
- $\ker N_{22} \subseteq \ker N_{12}$.

Then,

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top M \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0 \text{ for all } Z \in \mathbb{R}^{r \times q} \text{ such that } \begin{bmatrix} I \\ Z \end{bmatrix}^\top N \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0$$

if and only if there exists a scalar $\alpha \geqslant 0$ such that $M - \alpha N \geqslant 0$.

Setup noisy input-state-output data

data-consistent systems

$$\Sigma = \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} : \begin{bmatrix} I \\ \hline A^\top \; C^\top \\ B^\top \; D^\top \end{bmatrix}^\top \underbrace{\left[\begin{array}{c|c} I & X_+ \\ & Y_- \\ \hline 0 & -X_- \\ & -U_- \end{array}\right] \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \left[\begin{array}{c|c} I & X_+ \\ & Y_- \\ \hline 0 & -X_- \\ & -U_- \end{array}\right]^\top}_{N} \begin{bmatrix} I \\ \hline A^\top \; C^\top \\ B^\top \; D^\top \end{bmatrix} \geqslant 0 \right\}$$

Recall dissipativity QMI

$$\begin{bmatrix} I \\ \hline A^\top \; C^\top \\ B^\top \; D^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} P^{-1} & 0 & 0 & 0 \\ 0 & R_{yy} & 0 & R_{yu} \\ 0 & 0 & -P^{-1} & 0 \\ 0 & R_{uy} & 0 & R_{uu} \end{bmatrix}}_{M} \begin{bmatrix} I \\ \hline A^\top \; C^\top \\ B^\top \; D^\top \end{bmatrix} \geqslant 0$$

Theorem an LMI test

Suppose that

- $S \in \mathbb{S}^{m+p}$ has exactly p negative and m positive eigenvalues. (for dualization)
- $N_{22} \leqslant 0$, $N_{11} - N_{12}N_{22}^{-1}N_{21} \gneqq 0$, and $\ker N_{22} \subseteq \ker N_{12}$.

Then, (U_-, X, Y_-) are informative for dissipativity w.r.t. the supply rate s if and only if there exist $Q > 0$ and a positive number α s.t.

$$\begin{bmatrix} Q & 0 & 0 & 0 \\ 0 & R_{yy} & 0 & R_{yu} \\ 0 & 0 & -Q & 0 \\ 0 & R_{uy} & 0 & R_{uu} \end{bmatrix} - \alpha \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right] \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \left[\begin{array}{c|c} I & \begin{matrix} X_+ \\ Y_- \end{matrix} \\ \hline 0 & \begin{matrix} -X_- \\ -U_- \end{matrix} \end{array}\right]^\top \geqslant 0$$

where

$$\begin{bmatrix} R_{yy} & R_{yu} \\ R_{uy} & R_{uu} \end{bmatrix} = \begin{bmatrix} 0 & -I_p \\ I_m & 0 \end{bmatrix} \begin{bmatrix} S_{uu} & S_{uy} \\ S_{yu} & S_{yy} \end{bmatrix}^{-1} \begin{bmatrix} 0 & -I_m \\ I_p & 0 \end{bmatrix}.$$

Recap

$\mathcal{H}_\infty$ control

Dissipativity analysis

Stabilization from input-output data

Summary and outlook

Discussion

auto-regressive (AR) models

- **To-be-controlled system:** Consider an IO system of the form

$$y(t+L) + P_{L-1}y(t+L-1) + \cdots + P_0 y(t) = Q_{L-1}u(t+L-1) + \cdots + Q_0 u(t)$$

where $y \in \mathbb{R}^p$ is the output, $u \in \mathbb{R}^m$ is the input, $P_i \in \mathbb{R}^{p\times p}$, and $Q_i \in \mathbb{R}^{p\times m}$.

- $L > 0$: order of the system

- **Notation:** $P(\sigma)y = Q(\sigma)u$ where σ is the shift operator $(\sigma \boldsymbol{f})(t) = \boldsymbol{f}(t+1)$ and

$$P(\xi) = I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_1\xi + P_0,$$
$$Q(\xi) = 0\xi^L + Q_{L-1}\xi^{L-1} + \cdots + Q_1\xi + Q_0.$$

- **Autonomous systems ($\equiv$ no inputs):** $P(\sigma)y = \mathbf{0}$

Definition stability

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if $\lim_{t\to\infty} y(t) = 0$ for all trajectories $y : \mathbb{Z}_+ \to \mathbb{R}^p$.

Theorem algebraic characterization of stability

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if and only if all roots of $\det(P(\xi))$ are in $\{\lambda \in \mathbb{C} : |\lambda| < 1\}$.

Reminder

$P(\xi) = I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_1\xi + P_0$

Theorem Lyapunov characterization

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if and only if there exists $\Psi \in \mathbb{S}^{pL}$ such that $\Psi \geqslant 0$ and

$$\begin{bmatrix} I_{pL} \\ -\boldsymbol{P} \end{bmatrix}^\top \left(\begin{bmatrix} \Psi & 0 \\ 0 & 0_{p,p} \end{bmatrix} - \begin{bmatrix} 0_{p,p} & 0 \\ 0 & \Psi \end{bmatrix} \right) \begin{bmatrix} I_{pL} \\ -\boldsymbol{P} \end{bmatrix} > 0$$

where $\boldsymbol{P} = \begin{bmatrix} P_0 & P_1 & \cdots & P_{L-1} \end{bmatrix}$.

Remark

Proof is based on behavioral theory, in particular quadratic difference forms.

Theorem — Lyapunov characterization

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if and only if there exists $\Psi \in \mathbb{S}^{pL}$ such that $\Psi \geqslant 0$ and

$$\begin{bmatrix} I_{pL} \\ -\boldsymbol{P} \end{bmatrix}^\top \left(\begin{bmatrix} \Psi & 0 \\ 0 & 0_{p,p} \end{bmatrix} - \begin{bmatrix} 0_{p,p} & 0 \\ 0 & \Psi \end{bmatrix} \right) \begin{bmatrix} I_{pL} \\ -\boldsymbol{P} \end{bmatrix} > 0 \qquad (\star)$$

where $\boldsymbol{P} = \begin{bmatrix} P_0 & P_1 & \cdots & P_{L-1} \end{bmatrix}$.

Observation

$$(\star) \iff \left[\begin{array}{cc} I_p & 0_{p,p(L-1)} \\ 0_{p(L-1)\times p} & I_{p(L-1)} \\ \hline \multicolumn{2}{c}{-\boldsymbol{P}} \end{array}\right]^\top \left(\begin{bmatrix} \Psi & 0 \\ 0 & 0_{p,p} \end{bmatrix} - \begin{bmatrix} 0_{p,p} & 0 \\ 0 & \Psi \end{bmatrix} \right) \left[\begin{array}{cc} I_p & 0_{p,p(L-1)} \\ 0_{p(L-1)\times p} & I_{p(L-1)} \\ \hline \multicolumn{2}{c}{-\boldsymbol{P}} \end{array}\right] > 0$$

$$\iff \Psi - \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix} > 0 \implies \Psi > 0 \qquad \text{where } J := \begin{bmatrix} 0_{p(L-1)\times p} & I_{p(L-1)} \end{bmatrix}.$$

Theorem — refinement of Lyapunov characterization

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if and only if there exists $\Psi \in \mathbb{S}^{pL}$ such that $\Psi > 0$ and $\Psi - \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix} > 0$ where $J = \begin{bmatrix} 0_{p(L-1)\times p} & I_{p(L-1)} \end{bmatrix}$ and $\boldsymbol{P} = \begin{bmatrix} P_0 & P_1 & \cdots & P_{L-1} \end{bmatrix}$.

Discussion auto-regressive (AR) models

- **To-be-controlled system:** Consider an IO system of the form

$$y(t+L) + P_{L-1}y(t+L-1) + \cdots + P_0 y(t) = Q_{L-1}u(t+L-1) + \cdots + Q_0 u(t)$$

where $y \in \mathbb{R}^p$ is the output, $u \in \mathbb{R}^m$ is the input, $P_i \in \mathbb{R}^{p\times p}$, and $Q_i \in \mathbb{R}^{p\times m}$.

- $L > 0$: order of the system

- **Notation:** $P(\sigma)y = Q(\sigma)u$ where σ is the shift operator $(\sigma \boldsymbol{f})(t) = \boldsymbol{f}(t+1)$ and

$$P(\xi) = I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_1\xi + P_0,$$

$$Q(\xi) = 0\xi^L + Q_{L-1}\xi^{L-1} + \cdots + Q_1\xi + Q_0.$$

- **Autonomous systems (≡ no inputs):** $P(\sigma)y = \mathbf{0}$

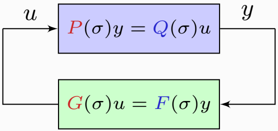
- **Controller:** Consider $G(\sigma)u = F(\sigma)y$ where u is the **output**, y is the **input**,

$$G(\xi) = I\xi^L + G_{L-1}\xi^{L-1} + \cdots + G_1\xi + G_0,$$

$$F(\xi) = 0\xi^L + F_{L-1}\xi^{L-1} + \cdots + F_1\xi + F_0.$$

- **Closed-loop:**

$$\begin{bmatrix} G(\sigma) & -F(\sigma) \\ -Q(\sigma) & P(\sigma) \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \mathbf{0}$$



Recall

- **To-be-controlled system:** $P(\sigma)y = Q(\sigma)u$ where

$$P(\xi) = I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_1\xi + P_0 \quad \text{and} \quad Q(\xi) = 0\xi^L + Q_{L-1}\xi^{L-1} + \cdots + Q_1\xi + Q_0.$$

- **Controller:** $G(\sigma)u = F(\sigma)y$ where

$$G(\xi) = I\xi^L + G_{L-1}\xi^{L-1} + \cdots + G_1\xi + G_0 \quad \text{and} \quad F(\xi) = 0\xi^L + F_{L-1}\xi^{L-1} + \cdots + F_1\xi + F_0.$$

- **Closed-loop:** $\begin{bmatrix} G(\sigma) & -F(\sigma) \\ -Q(\sigma) & P(\sigma) \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \mathbf{0}$ where

$$\begin{bmatrix} G(\xi) & -F(\xi) \\ -Q(\xi) & P(\xi) \end{bmatrix} = I\xi^L + \begin{bmatrix} G_{L-1} & -F_{L-1} \\ -Q_{L-1} & P_{L-1} \end{bmatrix} \xi^{L-1} + \cdots + \begin{bmatrix} G_0 & -F_0 \\ -Q_0 & P_0 \end{bmatrix}$$

Recall

Closed-loop: $\begin{bmatrix} G(\sigma) & -F(\sigma) \\ -Q(\sigma) & P(\sigma) \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \mathbf{0}$ where

$$\begin{bmatrix} G(\xi) & -F(\xi) \\ -Q(\xi) & P(\xi) \end{bmatrix} = I\xi^L + \begin{bmatrix} G_{L-1} & -F_{L-1} \\ -Q_{L-1} & P_{L-1} \end{bmatrix} \xi^{L-1} + \cdots + \begin{bmatrix} G_0 & -F_0 \\ -Q_0 & P_0 \end{bmatrix}$$

Theorem

refinement of Lyapunov characterization

An autonomous system of the form $P(\sigma)y = \mathbf{0}$ is stable if and only if there exists $\Psi \in \mathbb{S}^{pL}$ such that $\Psi > 0$ and $\Psi - \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{P} \end{bmatrix} > 0$ where $J = \begin{bmatrix} 0_{p(L-1)\times p} & I_{p(L-1)} \end{bmatrix}$ and $\boldsymbol{P} = \begin{bmatrix} P_0 & P_1 & \cdots & P_{L-1} \end{bmatrix}$.

Notation

$\boldsymbol{C} := \begin{bmatrix} G_0 & -F_0 & G_1 & -F_1 & \cdots & G_{L-1} & -F_{L-1} \end{bmatrix}$ and $\boldsymbol{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$.

Theorem

single system stabilization

The closed loop system is stable if and only if there exists $\Psi \in \mathbb{S}^{(m+p)L}$ such that $\Psi > 0$ and $\Psi - \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix} > 0$ where $J = \begin{bmatrix} 0_{(m+p)(L-1)\times p} & I_{(m+p)(L-1)} \end{bmatrix}$.

Setup

input-output data

model class $P(\sigma)y = Q(\sigma)u + w$ $\qquad y \in \mathbb{R}^p, u \in \mathbb{R}^m, w \in \mathbb{R}^p$

$$P(\xi) = I\xi^L + P_{L-1}\xi^{L-1} + \cdots + P_0 \quad \text{and} \quad Q(\xi) = 0\xi^L + Q_{L-1}\xi^{L-1} + \cdots + Q_0.$$

$$y(t+L) + P_{L-1}y(t+L-1) + \cdots + P_0 y(t) - Q_{L-1}u(t+L-1) - \cdots - Q_0 u(t) = w(t)$$

true system $P_s(\sigma)y = Q_s(\sigma)u + w$ $\qquad P_s, Q_s, w$ are unknown

data $U_- = \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix} \qquad Y_- = \begin{bmatrix} y(0) & y(1) & \cdots & y(T-1) \end{bmatrix}$

observe

$$\boldsymbol{v}(t) = \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} \Longrightarrow \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} & 0 & I \end{bmatrix} \begin{bmatrix} \boldsymbol{v}(t) \\ \boldsymbol{v}(t+1) \\ \vdots \\ \boldsymbol{v}(t+L-1) \\ \boldsymbol{v}(t+L) \end{bmatrix} = w(t)$$

recall $\boldsymbol{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$

data-consistent systems $\qquad \boldsymbol{R}$ is a data-consistent system if and only if

$$\begin{bmatrix} \boldsymbol{R} & 0 & I \end{bmatrix} \begin{bmatrix} v(0) & v(1) & \cdots & v(T-L-1) \\ v(1) & v(2) & \cdots & v(T-L) \\ \vdots & \vdots & & \vdots \\ v(L) & v(L+1) & \cdots & v(T-1) \end{bmatrix} = \overbrace{\begin{bmatrix} w(0) & w(1) & \cdots & w(T-L-1) \end{bmatrix}}^{W_-}$$

Setup

input-output data

data-consistent systems $\boldsymbol{R}$ is a data-consistent system if and only if

$$\begin{bmatrix} \boldsymbol{R} & I \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} = W_-$$

noise model $W_- = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L-1) \end{bmatrix}$ is unknown but satisfies:

$$\begin{bmatrix} I \\ W_-^\top \end{bmatrix}^\top \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I \\ W_-^\top \end{bmatrix} \geqslant 0 \qquad (\Phi_{11} \in \mathbb{S}^p \quad \text{and} \quad \Phi_{22} \in \mathbb{S}^{T-L})$$

data-consistent systems $\boldsymbol{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$

$$\Sigma = \left\{ \boldsymbol{R} : \begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix}^\top \underbrace{\begin{bmatrix} I & H_2 \\ 0 & H_1 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \begin{bmatrix} I & H_2 \\ 0 & H_1 \end{bmatrix}^\top}_{\boldsymbol{N}} \begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix} \geqslant 0 \right\}$$

Setup input-output data

data-consistent systems $\boldsymbol{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$

$$\Sigma = \left\{ \boldsymbol{R} : \begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix}^\top \boldsymbol{N} \begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix} \geqslant 0 \right\}$$

Notation

$\boldsymbol{C} := \begin{bmatrix} G_0 & -F_0 & G_1 & -F_1 & \cdots & G_{L-1} & -F_{L-1} \end{bmatrix}$ and $\boldsymbol{R} := \begin{bmatrix} -Q_0 & P_0 & -Q_1 & P_1 & \cdots & -Q_{L-1} & P_{L-1} \end{bmatrix}$.

Theorem single system stabilization

The closed loop system is stable if and only if there exists $\Psi \in \mathbb{S}^{(m+p)L}$ such that $\Psi > 0$ and $\Psi - \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix} > 0$ where $J = \begin{bmatrix} 0_{(m+p)(L-1)\times p} & I_{(m+p)(L-1)} \end{bmatrix}$.

Definition stabilization by dynamic feedback

(U_-, Y_-) are informative for quadratic stabilization if there exists $\boldsymbol{C} \in \mathbb{R}^{m\times(m+p)L}$ and $\Psi \in \mathbb{S}^{(m+p)L}$ such that $\Psi > 0$ and $\Psi - \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -\boldsymbol{C} \\ -\boldsymbol{R} \end{bmatrix} > 0$ for all $\boldsymbol{R} \in \Sigma$ where $J = \begin{bmatrix} 0_{(m+p)(L-1)\times p} & I_{(m+p)(L-1)} \end{bmatrix}$.

Observation stabilization by dynamic feedback

$$\Psi - \begin{bmatrix} J \\ -C \\ -R \end{bmatrix}^\top \Psi \begin{bmatrix} J \\ -C \\ -R \end{bmatrix} > 0 \iff \begin{bmatrix} \Psi & \begin{bmatrix} J \\ -C \\ -R \end{bmatrix}^\top \\ \begin{bmatrix} J \\ -C \\ -R \end{bmatrix} & \Psi^{-1} \end{bmatrix} > 0$$

$$\iff \Psi^{-1} - \begin{bmatrix} J \\ -C \\ -R \end{bmatrix} \Psi^{-1} \begin{bmatrix} J \\ -C \\ -R \end{bmatrix}^\top > 0$$

$$\overset{\Omega=\Psi^{-1}}{\iff} \begin{bmatrix} & I_{qL} \\ R^\top & \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix}^\top \underbrace{\begin{bmatrix} \Omega - \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix} \Omega \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix}^\top & \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix} \Omega \\ \Omega \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix}^\top & -\Omega \end{bmatrix}}_{M} \begin{bmatrix} & I_{qL} \\ R^\top & \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix} > 0$$

Question

When does the QMI

$$\begin{bmatrix} I_{qL} \\ R^\top \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix}^\top M \begin{bmatrix} I_{qL} \\ R^\top \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix} > 0$$

hold for all R such that

$$\begin{bmatrix} I \\ R^\top \end{bmatrix}^\top N \begin{bmatrix} I \\ R^\top \end{bmatrix} \geqslant 0 \quad ?$$

Theorem projection

Let $\Pi \in \mathbb{S}^{q+r}$ be partitioned as $\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix}$ where $\Pi_{11} \in \mathbb{S}^q$ and satisfy $\Pi_{22} < 0$ and $\Pi_{11} - \Pi_{12}\Pi_{22}^{-1}\Pi_{21} \geqslant 0$. Also, let $Y \in \mathbb{R}^{r\times p}$ and $W \in \mathbb{R}^{q\times p}$. Then, there exists a $Z \in \mathbb{R}^{r\times q}$ such that

$$\begin{bmatrix} I \\ Z \end{bmatrix}^\top \Pi \begin{bmatrix} I \\ Z \end{bmatrix} \geqslant 0 \quad \text{and} \quad ZW = Y$$

if and only if

$$\begin{bmatrix} I \\ Y \end{bmatrix}^\top \begin{bmatrix} W^\top \Pi_{11} W & W^\top \Pi_{12} \\ \Pi_{21} W & \Pi_{22} \end{bmatrix} \begin{bmatrix} I \\ Y \end{bmatrix} \geqslant 0.$$

Observation

Suppose that $N_{22} = H_1 \Phi_{22} H_1^\top < 0$. Then,

$$\begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix}^\top \boldsymbol{N} \begin{bmatrix} I \\ \boldsymbol{R}^\top \end{bmatrix} \geqslant 0$$

if and only if

$$\begin{bmatrix} & I_{qL} \\ \boldsymbol{R}^\top \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix}^\top \underbrace{\begin{bmatrix} \begin{bmatrix} 0 & -I_p \end{bmatrix} & 0 \\ 0 & I_{qL} \end{bmatrix}^\top \boldsymbol{N} \begin{bmatrix} \begin{bmatrix} 0 & -I_p \end{bmatrix} & 0 \\ 0 & I_{qL} \end{bmatrix}}_{\hat{\boldsymbol{N}}} \begin{bmatrix} & I_{qL} \\ \boldsymbol{R}^\top \begin{bmatrix} 0 & -I_p \end{bmatrix} \end{bmatrix} \geqslant 0.$$

Theorem

(U_-, Y_-) are informative for quadratic stabilization if and only if there exist $C \in \mathbb{R}^{m\times(m+p)L}$ and $\Omega \in \mathbb{S}^{(m+p)L}$ such that $\Omega > 0$ and

$$\begin{bmatrix} \Omega - \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix} \Omega \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix}^\top & \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix} \Omega \\ \Omega \begin{bmatrix} J \\ -C \\ 0 \end{bmatrix}^\top & -\Omega \end{bmatrix} - \hat{N} > 0$$

where $J = \begin{bmatrix} 0_{(m+p)(L-1)\times p} & I_{(m+p)(L-1)} \end{bmatrix}$.

Remark

The above matrix inequality is not an LMI. As it is done before for stabilization by state feedback, we can convert it to an LMI.

Recap

$\mathcal{H}_\infty$ control

Dissipativity analysis

Stabilization from input-output data

Summary and outlook

Exact data

Problem	Data
controllability	IS
observability	SO
stabilizability	IS
stability	S
stabilization by state feedback	IS
deadbeat controller	IS
LQR	IS
suboptimal LQR	IS
suboptimal $\mathcal{H}_2$	IS
stabilization by dynamic feedback	I(S)O
dissipativity analysis	I(S)O
tracking and regulation	IS
model reduction (moment matching)	IO

Noisy data

Problem	Data
stability	S
controllability	IS
stabilizability	IS
stabilization by state feedback	IS
$\mathcal{H}_2$ optimal control	IS
$\mathcal{H}_\infty$ control	IS
stability	IO
dissipativity analysis	I(S)O
model reduction (balancing)	ISO
structural properties	ISO
stabilization by dynamic feedback	IO

Input **S**tate **O**utput

Question. How can we device *online experiments* to obtain informative data?

Exact data

Problem	Data
controllability	IS
observability	SO
stabilizability	IS
stability	S
stabilization by state feedback	IS
deadbeat controller	IS
LQR	IS
suboptimal LQR	IS
suboptimal $\mathcal{H}_2$	IS
stabilization by dynamic feedback	I(S)O
dissipativity analysis	I(S)O
tracking and regulation	IS
model reduction (moment matching)	IO

Noisy data

Problem	Data
stability	S
controllability	IS
stabilizability	IS
stabilization by state feedback	IS
$\mathcal{H}_2$ optimal control	IS
$\mathcal{H}_\infty$ control	IS
stability	IO
dissipativity analysis	I(S)O
model reduction (balancing)	ISO
structural properties	ISO
stabilization by dynamic feedback	IO

Input **S**tate **O**utput

Question. How can we device *online experiments* to obtain informative data?

Data-based systems and control theory

Day 5

GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and Henk van Waarde

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen

Integer intervals:

- For $i \leqslant j$, $[i, j]$= the ordered set of integers from i to j
- For $i > j$, $[i, j]$= $\varnothing$

Matrices from ordered set of vectors:

Given vectors $f_0, f_1, \ldots, f_k$, $f_{[0,k]} = \begin{bmatrix} f_0 & f_1 & \cdots & f_k \end{bmatrix}$

Void matrices: Matrices with zero rows and/or columns.

The lag of an input-state-output system:

$$x(t+1) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$
$$u \in \mathbb{R}^m,\ x \in \mathbb{R}^n,\ y \in \mathbb{R}^p$$
$$m \geqslant 1,\ n \geqslant 0,\ p \geqslant 1$$

$$\Omega_k = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}$$

The lag is defined as the smallest integer $\ell \geqslant 0$ such that $\operatorname{rank} \Omega_\ell = \operatorname{rank} \Omega_{\ell+1}$ and denoted by $\ell(C, A)$. $0 \leqslant \ell \leqslant n$

True system:

$$x(t+1) = A_{\text{true}}x(t) + B_{\text{true}}u(t)$$
$$y(t) = C_{\text{true}}x(t) + D_{\text{true}}u(t)$$

$n_{\text{true}}, \ell_{\text{true}}, A_{\text{true}}, B_{\text{true}}, C_{\text{true}}, D_{\text{true}}$ **unknown**

$(A_{\text{true}}, B_{\text{true}}, C_{\text{true}})$ **minimal**

$n_{\text{true}} \geqslant 0$

$\ell_{\text{true}} \geqslant 0$

$A_{\text{true}} \in \mathbb{R}^{n_{\text{true}} \times n_{\text{true}}}$

$B_{\text{true}} \in \mathbb{R}^{n_{\text{true}} \times m}$

$C_{\text{true}} \in \mathbb{R}^{p \times n_{\text{true}}}$

$D_{\text{true}} \in \mathbb{R}^{p \times m}$

Data: Let $T \geqslant 1$ and $(u_{[0,T-1]}, y_{[0,T-1]})$ be generated by the true system, i.e.

$$\begin{bmatrix} \boldsymbol{x_{[1,T]}} \\ y_{[0,T-1]} \end{bmatrix} = \begin{bmatrix} A_{\text{true}} & B_{\text{true}} \\ C_{\text{true}} & D_{\text{true}} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_{[0,T-1]}} \\ u_{[0,T-1]} \end{bmatrix}$$

for some $\boldsymbol{x_{[0,T]}} \in \mathbb{R}^{n_{\text{true}} \times (T+1)}$.

Question. Under what conditions can we identify (up to a state space transformation) the true system?

Obs. Not only data but also prior knowledge plays a role.

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \quad \longleftrightarrow \quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{(n+\boldsymbol{p})\times(n+\boldsymbol{m})}$$

$\mathcal{S} = \{[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathbb{R}^{(n+\boldsymbol{p})\times(n+\boldsymbol{m})} : n \geqslant 0\}$ ISO systems with $\boldsymbol{m}$ inputs and $\boldsymbol{p}$ outputs

$\mathcal{O} = \{[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathcal{S} : (C, A) \text{ is observable}\}$ observable systems

$\mathcal{M} = \{[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathcal{O} : (A, B) \text{ is controllable}\}$ minimal systems

---

$\mathcal{S}(\boldsymbol{n}) = \{[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathcal{S} : A \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}}\}$ systems with $\boldsymbol{n}$ states

$\mathcal{S}(\boldsymbol{\ell}, \boldsymbol{n}) = \{[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathcal{S}(n) : \ell(C, A) = \ell\}$ systems with lag $\boldsymbol{\ell}$ and $\boldsymbol{n}$ states

Def. A system $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathbb{R}^{(n+\boldsymbol{p})\times(n+\boldsymbol{m})}$ explains the data $(u_{[0,T-1]}, y_{[0,T-1]})$ if

$$\begin{bmatrix} \boldsymbol{x_{[1,T]}} \\ y_{[0,T-1]} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \boldsymbol{x_{[0,T-1]}} \\ u_{[0,T-1]} \end{bmatrix}$$

for some $\boldsymbol{x_{[0,T]}} \in \mathbb{R}^{n\times(T+1)}$.

$\mathcal{E} = \left\{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S} : \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \text{ explains the data } (u_{[0,T-1]}, y_{[0,T-1]}) \right\}$ explaining systems

$\mathcal{E}(\boldsymbol{n}) = \mathcal{E} \cap \mathcal{S}(n)$ explaining systems with $\boldsymbol{n}$ states

$\mathcal{E}(\boldsymbol{\ell}, \boldsymbol{n}) = \mathcal{E} \cap \mathcal{S}(\ell, n)$ explaining systems with lag $\boldsymbol{\ell}$ and $\boldsymbol{n}$ states

---

True system: $\begin{bmatrix} A_{\text{true}} & B_{\text{true}} \\ C_{\text{true}} & D_{\text{true}} \end{bmatrix} \in \mathcal{E}(\ell_{\text{true}}, n_{\text{true}}) \subseteq \mathcal{E}(n_{\text{true}}) \subseteq \mathcal{E}$

Def. Two systems $\begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \in \mathcal{S}(n)$ are isomorphic if $D_1 = D_2$ and $\exists$ a nonsingular matrix $S \in \mathbb{R}^{n \times n}$ such that $A_1 = S^{-1} A_2 S$, $B_1 = S^{-1} B_2$, $C_1 = C_2 S$.

Def. A set of systems $\mathcal{S}' \subseteq \mathcal{S}(n)$ has the isomorphism property if all systems belonging to $\mathcal{S}'$ are isomorphic to each other.

Prior knowledge: $\boldsymbol{\mathcal{S}_{\mathrm{pk}}} \subseteq \mathcal{S}$ with $\begin{bmatrix} A_{\mathrm{true}} & B_{\mathrm{true}} \\ C_{\mathrm{true}} & D_{\mathrm{true}} \end{bmatrix} \in \mathcal{S}_{\mathrm{pk}}$

Def. The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysID within $\mathcal{S}_{\mathrm{pk}}$ if

- $\mathcal{E} \cap \mathcal{S}_{\mathrm{pk}} = \mathcal{E}(n_{\mathrm{true}}) \cap \mathcal{S}_{\mathrm{pk}}$ (data determine state dimension)
- $\mathcal{E}(n_{\mathrm{true}}) \cap \mathcal{S}_{\mathrm{pk}}$ has the isomorphism property

- We assume that upper bounds (L, N) on the true lag and state dimension are given:

$$\ell_{\text{true}} \leqslant L \quad \text{and} \quad n_{\text{true}} \leqslant N.$$

- We define

$$\mathcal{S}_{L,N} := \bigcup_{\substack{\ell \in [0, L] \\ n \in [0, N]}} \mathcal{S}(\ell, n).$$

- We are interested in system identification within

$$\boldsymbol{\mathcal{S}_{\text{pk}} = \mathcal{S}_{L,N} \cap \mathcal{M}}.$$

Prop. fundamental lemma

Suppose that

$$T \geqslant L + (L+N+1)m + N.$$

If the input $u_{[0,T-1]}$ is persistently exciting of order $N+L+1$, i.e.

$$\operatorname{rank} H_{L+N+1}(\boxed{u_{[0,T-1]}}) = (L+N+1)m,$$

then the data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysID within $\mathcal{S}_{L,N} \cap \mathcal{M}$.

Remarks. It is a

- sufficient condition that is not necessary.
- condition only on the input.

$\ell_{\min} = \min \{\ell \geqslant 0 : \exists\, n \geqslant 0 \text{ s.t. } \mathcal{E}(\ell, n) \neq \varnothing\}$ minimum lag to explain the data

$n_{\min} = \min \{n \geqslant 0 : \mathcal{E}(n) \neq \varnothing\}$ minimum state dimension to explain the data

---

Question. How can we obtain $\ell_{\min}$ and $n_{\min}$ from the data?

$$H_k = \left[\begin{array}{ccc} y(0) & \cdots & y(T-k) \\ y(1) & \cdots & y(T-k+1) \\ \vdots & & \vdots \\ y(k-2) & \cdots & y(T-2) \\ y(k-1) & \cdots & y(T-1) \\ \hline u(0) & \cdots & u(T-k) \\ u(1) & \cdots & u(T-k+1) \\ \vdots & & \vdots \\ u(k-2) & \cdots & u(T-2) \\ u(k-1) & \cdots & u(T-1) \end{array}\right] \text{ and } G_k = \left[\begin{array}{ccc} y(0) & \cdots & y(T-k) \\ y(1) & \cdots & y(T-k+1) \\ \vdots & & \vdots \\ y(k-2) & \cdots & y(T-2) \\ \hline u(0) & \cdots & u(T-k) \\ u(1) & \cdots & u(T-k+1) \\ \vdots & & \vdots \\ u(k-2) & \cdots & u(T-2) \\ u(k-1) & \cdots & u(T-1) \end{array}\right]$$

$\ell_{\min} = \min\{\ell \geqslant 0 : \exists\, n \geqslant 0 \text{ s.t. } \mathcal{E}(\ell, n) \neq \varnothing\}$ minimum lag to explain the data

$n_{\min} = \min\{n \geqslant 0 : \mathcal{E}(n) \neq \varnothing\}$ minimum state dimension to explain the data

---

Question. How can we obtain $\ell_{\min}$ and $n_{\min}$ from the data?

Assumption. $u_{[0,T-1]} \neq 0_{m,T}$ (necessary for SysID)

$$\delta_k = \begin{cases} p & \text{if} \quad k = 0 \\ \operatorname{rank} H_k - \operatorname{rank} G_k & \text{if} \quad k \in [1, T] \end{cases}$$

$$p = \delta_0 \geqslant \delta_1 \geqslant \cdots \geqslant \delta_{T-1} \geqslant \delta_T = 0$$

$q := \min\{k \in [0, T-1] : \delta_{k+1} = 0\}$ $\quad q \in [0, T-1]$

Thm. Suppose that $\mathcal{E}(\ell, n) \neq \varnothing$. Then, the following statements hold:

- If $T \geqslant \ell + 1$, then $\ell \geqslant q$.
- If $\ell \geqslant q$, then $n - \sum_{i=1}^{q} \delta_i \geqslant \ell - q$.

$\ell_{\min} = \min\{\ell \geqslant 0 : \exists\, n \geqslant 0 \text{ s.t. } \mathcal{E}(\ell, n) \neq \varnothing\}$ minimum lag to explain the data

$n_{\min} = \min\{n \geqslant 0 : \mathcal{E}(n) \neq \varnothing\}$ minimum state dimension to explain the data

---

Thm. Suppose that $\mathcal{E}(\ell, n) \neq \varnothing$. Then, the following statements hold:

- If $T \geqslant \ell + 1$, then $\ell \geqslant q$.
- If $\ell \geqslant q$, then $n - \sum_{i=1}^{q} \delta_i \geqslant \ell - q$.

Def. We say that $\boldsymbol{x}_{[0,T]} \in \mathbb{R}^{n \times (T+1)}$ is a state for the data $(u_{[0,T-1]}, y_{[0,T-1]})$ if there exists $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S}(n)$ such that $\begin{bmatrix} \boldsymbol{x}_{[1,T]} \\ y_{[0,T-1]} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{[0,T-1]} \\ u_{[0,T-1]} \end{bmatrix}$.

Lem. $\boldsymbol{x}_{[0,T]} \in \mathbb{R}^{n \times (T+1)}$ is a state for the data $(u_{[0,T-1]}, y_{[0,T-1]})$ if and only if

$$\operatorname{rsp} \begin{bmatrix} \boldsymbol{x}_{[1,T]} \\ y_{[0,T-1]} \end{bmatrix} \subseteq \operatorname{rsp} \begin{bmatrix} \boldsymbol{x}_{[0,T-1]} \\ u_{[0,T-1]} \end{bmatrix}.$$

Thm. $\varnothing \neq \mathcal{E}(\sum_{i=1}^{q} \delta_i) = \mathcal{E}(q, \sum_{i=1}^{q} \delta_i)$.

Proof is based on a novel state construction.

$\ell_{\min} = \min\{\ell \geqslant 0 : \exists\, n \geqslant 0 \text{ s.t. } \mathcal{E}(\ell, n) \neq \varnothing\}$ minimum lag to explain the data

$n_{\min} = \min\{n \geqslant 0 : \mathcal{E}(n) \neq \varnothing\}$ minimum state dimension to explain the data

---

$q := \min\{k \in [0, T-1] : \delta_{k+1} = 0\}$ $q \in [0, T-1]$

Thm. Suppose that $\mathcal{E}(\ell, n) \neq \varnothing$. Then, the following statements hold:

- If $T \geqslant \ell + 1$, then $\ell \geqslant q$.
- If $\ell \geqslant q$, then $n - \sum_{i=1}^{q} \delta_i \geqslant \ell - q$.

Thm. $\varnothing \neq \mathcal{E}(\sum_{i=1}^{q} \delta_i) = \mathcal{E}(q, \sum_{i=1}^{q} \delta_i)$.

Thm. $\ell_{\min} = q$, $n_{\min} = \sum_{i=1}^{\ell_{\min}} \delta_i$, and $\mathcal{E}(n_{\min}) = \mathcal{E}(\ell_{\min}, n_{\min}) = \mathcal{E}(n_{\min}) \cap \mathcal{O}$.

Cor. Let $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{E}(\ell, n)$. Then, $\ell \leqslant n - n_{\min} + \ell_{\min}$.

Obs. $L_{\mathrm{d}} := N - n_{\min} + \ell_{\min}$ upper bound determined by the data

$L_{\mathrm{a}} := \min(L, L_{\mathrm{d}})$ actual upper bound

Recall. $L_{\mathrm{d}} := N - n_{\min} + \ell_{\min}$ and $L_{\mathrm{a}} := \min(L, L_{\mathrm{d}})$

Thm. [Camlibel and Rapisarda, "Beyond Willems' fundamental lemma: from finite time series to linear system"]

The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysID if and only if

$$T \geqslant L_{\mathrm{a}} + (L_{\mathrm{a}}+1)m + n_{\min}$$

and

$$\operatorname{rank} H_{L_{\mathrm{a}}+1} = (L_{\mathrm{a}}+1)m + n_{\min}.$$

Moreover, if these are satisfied, then

$$\ell_{\mathrm{true}} = \ell_{\min},$$
$$n_{\mathrm{true}} = n_{\min},$$
$$\mathcal{E} \cap \mathcal{S}_{L,N} \cap \mathcal{M} = \mathcal{E}(n_{\min}).$$

Special cases:

- L is known but N is unknown: take $N = pL$ $\quad L_{\mathrm{a}} = L$
- N is known but L is unknown: take $L = N$ $\quad L_{\mathrm{a}} = L_{\mathrm{d}}$

Prop. fundamental lemma

Suppose that

$$T \geqslant L + (L + N + 1)m + N.$$

If the input $u_{[0,T-1]}$ is persistently exciting of order $N + L + 1$, i.e.

$$\operatorname{rank} H_{L+N+1}(\boxed{u_{[0,T-1]}}) = (L + N + 1)m,$$

then the data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysID within $\mathcal{S}_{L,N} \cap \mathcal{M}$.

Recall. $L_\mathrm{d} := N - n_{\min} + \ell_{\min}$ and $L_\mathrm{a} := \min(L, L_\mathrm{d})$

Thm. The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysID **if and only if**

$$T \geqslant L_\mathrm{a} + (L_\mathrm{a} + 1)m + n_{\min}$$

and

$$\operatorname{rank} H_{L_\mathrm{a}+1} = (L_\mathrm{a} + 1)m + n_{\min}.$$

Consider the true system where $\boldsymbol{m} = 2$, $\boldsymbol{p} = 2$, $\boldsymbol{\ell}_{\textbf{true}} = 2$, $\boldsymbol{n}_{\textbf{true}} = 3$, and

$$\left[\begin{array}{c|c} A_{\text{true}} & B_{\text{true}} \\ \hline C_{\text{true}} & D_{\text{true}} \end{array}\right] = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

Consider the input-output data generated by the true system:

$$\left[\begin{array}{c} u_{[0,13]} \\ \hline y_{[0,13]} \end{array}\right] = \left[\begin{array}{cccccccccccccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ \hline 2 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 4 \\ 1 & 0 & 0 & 0 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 \end{array}\right].$$

For $(u_{[0,T-1]}, y_{[0,T-1]})$, we have

T	1	2	$[3,5]$	$[6,14]$
δ_{-1}	2	2	2	2
δ_0	0	1	2	2
δ_1		0	0	1
δ_2			0	0
$\ell_{\min}$	0	1	1	2
$n_{\min}$	0	1	2	3

				T values			
L	N	L_{d}	L_{a}	11	12	13	14
2	3	2	2	✓	✓	✓	✓
2	4	3	2	✓	✓	✓	✓
2	5	4	2	✓	✓	✓	✓
2	6	5	2	✓	✓	✓	✓
3	3	2	2	✓	✓	✓	✓
3	4	3	3				✓
3	5	4	3				✓
3	6	5	3				✓
4	4	3	3				✓

Here: ✓ means the corresponding data are informative

Note. For the case $L = N = 4$, 26 data points are needed to infer informativity via the fundamental lemma.

Data-based systems and control theory

Day 5.2

GIAN course at IIT Mandi, April 2025

Kanat Çamlıbel and **Henk van Waarde**

Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence

Jan C. Willems Center for Systems and Control

University of Groningen

1. The problem of experiment design
   - fundamental lemma
   - online experiment design
2. Informativity for system identification
3. The shortest experiment

The problem of experiment design

Informativity for system identification

The shortest experiment

True system:

$$x(t+1) = A_{\text{true}} x(t) + B_{\text{true}} u(t)$$
$$y(t) = C_{\text{true}} x(t) + D_{\text{true}} u(t)$$

$\begin{bmatrix} A_{\text{true}} & B_{\text{true}} \\ C_{\text{true}} & D_{\text{true}} \end{bmatrix} \in \mathbb{R}^{(n_{\text{true}}+p)\times(n_{\text{true}}+m)}$ and n_{true} are **unknown**

High-level goal: Design inputs $u_{[0,T-1]}$ such that we can **identify** the true system from the resulting input-output data $(u_{[0,T-1]}, y_{[0,T-1]})$.

Observability matrix and lag:
We define

$$\boldsymbol{\Omega_k} = \begin{bmatrix} C_{\text{true}} \\ C_{\text{true}} A_{\text{true}} \\ \vdots \\ C_{\text{true}} A_{\text{true}}^{k-1} \end{bmatrix}.$$

The **lag** is defined as the smallest integer $\ell \geqslant 0$ such that $\operatorname{rank} \Omega_\ell = \operatorname{rank} \Omega_{\ell+1}$ and denoted by $\ell_{\text{true}} = \ell(C_{\text{true}}, A_{\text{true}}) \leqslant n_{\text{true}}$.

True system:

$$\begin{aligned} x(t+1) &= A_{\rm true}x(t) + B_{\rm true}u(t) \\ y(t) &= C_{\rm true}x(t) + D_{\rm true}u(t) \end{aligned} \tag{1}$$

$\begin{bmatrix} A_{\rm true} & B_{\rm true} \\ C_{\rm true} & D_{\rm true} \end{bmatrix} \in \mathbb{R}^{(n_{\rm true}+p)\times(n_{\rm true}+m)}$ and $n_{\rm true}$ are **unknown**

Prior knowledge: (1) is **controllable** and **observable**, $\ell_{\rm true} \leqslant L$ and $n_{\rm true} \leqslant N$

Fundamental question: How to find $T \in \mathbb{N}$ and

$$u_{[0,T-1]} := \begin{bmatrix} u(0) & u(1) & \cdots & u(T-1) \end{bmatrix}$$

such that the resulting **data** $(u_{[0,T-1]}, y_{[0,T-1]})$ **enable system identification**?

I.e., such that we can **identify** $n_{\rm true}$ and matrices A, B, C and D satisfying

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{\rm true} & B_{\rm true} \\ C_{\rm true} & D_{\rm true} \end{bmatrix} \begin{bmatrix} S^{-1} & 0 \\ 0 & I \end{bmatrix} \quad \text{for some invertible } S$$

A partial answer

A note on persistency of excitation

Jan C. Willems[a], Paolo Rapisarda[b], Ivan Markovsky[a,*], Bart L.M. De Moor[a]

[a]ESAT, SCD/SISTA, K.U. Leuven, Kasteelpark Arenberg 10, B 3001 Leuven, Heverlee, Belgium
[b]Department of Mathematics, University of Maastricht, 6200 MD Maastricht, The Netherlands

Received 3 June 2004; accepted 7 September 2004
Available online 30 November 2004

Abstract

We prove that if a component of the response signal of a controllable linear time-invariant system is persistently exciting of sufficiently high order, then the windows of the signal span the full system behavior. This is then applied to obtain conditions

---

Definition: The input $u_{[0,T-1]}$ is called **persistently exciting** of order k if

$$\operatorname{rank} H_k(u_{[0,T-1]}) = \operatorname{rank} \begin{bmatrix} u(0) & u(1) & \cdots & u(T-k) \\ \vdots & \vdots & & \vdots \\ u(k-1) & u(k) & \cdots & u(T-1) \end{bmatrix} = km$$

Possible solution:

- Choose $T := (N + L + 1)m + N + L$
- Design $u_{[0,T-1]}$ to be persistently exciting of order $N + L + 1$
- Then by the **fundamental lemma**,

$$\operatorname{rank} \begin{bmatrix} H_{L+1}(u_{[0,T-1]}) \\ \hline H_{L+1}(y_{[0,T-1]}) \end{bmatrix} = \left[\begin{array}{ccc} u(0) & \cdots & u(T-L-1) \\ \vdots & & \vdots \\ u(L) & \cdots & u(T-1) \\ \hline y(0) & \cdots & y(T-L-1) \\ \vdots & & \vdots \\ y(L) & \cdots & y(T-1) \end{array}\right] = (L+1)m + n_{\text{true}}$$

- Apply **subspace identification** to obtain A, B, C and D

We will now consider a simple example...

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

Define $T = 9$ and $u_{[0,8]} := \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$ (PE of order 5)

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,8]}) \\ \hdashline H_3(y_{[0,8]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hdashline -1 & 0 & 0 & 1 & 1 & 2 & 3 \\ 2 & 2 & 1 & 1 & 2 & 3 & 7 \\ 0 & 0 & 1 & 1 & 2 & 3 & 5 \\ 2 & 1 & 1 & 2 & 3 & 7 & 9 \\ 0 & 1 & 1 & 2 & 3 & 5 & 9 \\ 1 & 1 & 2 & 3 & 7 & 9 & 14 \end{bmatrix} = 5 \quad \Longrightarrow \quad n_{\text{true}} = 2$$

Possible solution:

- Choose $T := (N+L+1)m + N + L$
- Design $u_{[0,T-1]}$ to be persistently exciting of order $N+L+1$
- Then by the **fundamental lemma**,

$$\operatorname{rank}\begin{bmatrix} H_{L+1}(u_{[0,T-1]}) \\ H_{L+1}(y_{[0,T-1]}) \end{bmatrix} = \begin{bmatrix} u(0) & \cdots & u(T-L-1) \\ \vdots & & \vdots \\ u(L) & \cdots & u(T-1) \\ y(0) & \cdots & y(T-L-1) \\ \vdots & & \vdots \\ y(L) & \cdots & y(T-1) \end{bmatrix} = (L+1)m + n_{\text{true}}$$

- Apply **subspace identification** to obtain A, B, C and D

Question: Is this the **smallest possible** T? **Answer**: **no!**

IEEE CONTROL SYSTEMS LETTERS, VOL. 6, 2022 319

Beyond Persistent Excitation: Online Experiment Design for Data-Driven Modeling and Control

Henk J. van Waarde

***Abstract*—This letter presents a new experiment design method for data-driven modeling and control. The idea is to select inputs *online* (using past input/output data), leading to desirable rank properties of data Hankel matrices. In**

rank property is important, since it guarantees that *all* trajectories of the system can be parameterized in terms of the measured trajectory. Essentially, the Hankel matrix of measured inputs and outputs serves as a non-parametric model of

---

Possible solution:

- Design the input $u(t)$ **online** based on $(u_{[0,t-1]}, y_{[0,t-1]})$
- T is not specified a priori, but procedure terminates after $T = (L+1)m + n_{\text{true}} + L$ steps
- Apply **subspace identification** to obtain A, B, C and D

We again consider an example...

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$, measure $y_{[0,2]} = \begin{bmatrix} -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$;

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,2]}) \\ \hdashline H_2(y_{[0,1]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \hdashline -1 \\ 2 \\ 0 \\ 0 \end{bmatrix} = 1$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$, measure $y_{[0,2]} = \begin{bmatrix} -1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$;

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,3]}) \\ \hdashline H_2(y_{[0,2]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & \boldsymbol{u(3)} \\ \hdashline -1 & 0 \\ 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = 2 \text{ for } \boldsymbol{u(3)} = 1$$

Measure $y(3) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

True system and initial state:

$$A_{\mathrm{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\mathrm{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\mathrm{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\mathrm{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$ and design the rest of the inputs online

---

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,4]}) \\ \hdashline H_2(y_{[0,3]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & u(4) \\ \hdashline -1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} = 3 \text{ for any } u(4)$$

Take $u(4) = 0$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$ and design the rest of the inputs online

---

$$\operatorname{rank} \left[\begin{array}{c} H_3(u_{[0,5]}) \\ \hdashline H_2(y_{[0,4]}) \end{array} \right] = \operatorname{rank} \left[\begin{array}{cccc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & u(5) \\ \hdashline -1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{array} \right] = 4 \text{ for any } u(5)$$

Take $u(5) = 0$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \; B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$ and design the rest of the inputs online

---

$$\operatorname{rank} \left[\begin{array}{c} H_3(u_{[0,6]}) \\ \hdashline H_2(y_{[0,5]}) \end{array} \right] = \operatorname{rank} \left[\begin{array}{ccccc} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & u(6) \\ \hdashline -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 \end{array} \right] = 5 \text{ for any } u(6)$$

So we take $u(6) = 0$.

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \; B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$ and design the rest of the inputs online

---

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,7]}) \\ \hdashline H_2(y_{[0,6]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & u(7) \\ \hdashline -1 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 & 2 \end{bmatrix} = 5 \neq 6 \text{ for any } u(7)$$

So we do not apply $u(7)$ and **stop the procedure**.

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Define $u_{[0,2]} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \neq 0$ and design the rest of the inputs online

---

It follows that

$$\operatorname{rank} \begin{bmatrix} H_3(u_{[0,6]}) \\ \hdashline H_3(y_{[0,6]}) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hdashline -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 & 2 \end{bmatrix} = (L+1)m + n_{\text{true}} = 5 \implies n_{\text{true}} = 2$$

Reduction # of samples: from $T = 9$ to $T = 7$

Another partial answer

IEEE CONTROL SYSTEMS LETTERS, VOL. 6, 2022 319

Beyond Persistent Excitation: Online Experiment Design for Data-Driven Modeling and Control

Henk J. van Waarde

Abstract—**This letter presents a new experiment design method for data-driven modeling and control. The idea is to select inputs *online* (using past input/output data), leading to desirable rank properties of data Hankel matrices. In**

rank property is important, since it guarantees that *all* trajectories of the system can be parameterized in terms of the measured trajectory. Essentially, the Hankel matrix of measured inputs and outputs serves as a non-parametric model of

---

Possible solution:

- Design the input $u(t)$ **online** based on $(u_{[0,t-1]}, y_{[0,t-1]})$
- T is not specified a priori, but procedure terminates after $T = (L+1)m + n_{\text{true}} + L$ steps
- Apply **subspace identification** to obtain A, B, C and D

Question: Is this the **smallest possible** T? **Answer**: **it's a secret!**

The problem of experiment design

Informativity for system identification

The shortest experiment

Beyond the fundamental lemma:
from finite time series to linear system

M. Kanat Camlibel[1] and Paolo Rapisarda[2]

[1]Bernoulli Institute, University of Groningen
[2]School of Electronics and Computer Science, University of Southampton

Abstract

We state necessary and sufficient conditions to uniquely identify (modulo state isomorphism) a linear time-invariant minimal input-state-output system from finite input-output data and upper- and lower bounds on lag and state space dimension.

Data: Let $(u_{[0,T-1]}, y_{[0,T-1]})$ be generated by the true system (no assumptions on the input for now!)

Question: Under what conditions on $(u_{[0,T-1]}, y_{[0,T-1]})$ can we **uniquely identify** the true system (up to state-space transformations)?

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \quad \longleftrightarrow \quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+m)}$$

$\mathcal{S} = \left\{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathbb{R}^{(n+\boldsymbol{p})\times(n+\boldsymbol{m})} : n \geqslant 0 \right\}$ systems with $\boldsymbol{m}$ inputs and $\boldsymbol{p}$ outputs

$\mathcal{O} = \{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S} : (C, A)$ is **observable**$\}$ observable systems

$\mathcal{M} = \{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{O} : (A, B)$ is **controllable**$\}$ minimal systems

---

$\mathcal{S}(\boldsymbol{n}) = \{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S} : A \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{n}} \}$ systems with $\boldsymbol{n}$ states

$\mathcal{S}(\boldsymbol{\ell}, \boldsymbol{n}) = \{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S}(n) : \ell(C, A) = \boldsymbol{\ell} \}$ systems with lag $\boldsymbol{\ell}$ and $\boldsymbol{n}$ states

Definition: A system $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathbb{R}^{(n+p)\times(n+m)}$ explains the data $(u_{[0,T-1]}, y_{[0,T-1]})$ if

$$\begin{bmatrix} \boldsymbol{x}_{[1,T]} \\ y_{[0,T-1]} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{[0,T-1]} \\ u_{[0,T-1]} \end{bmatrix}$$

for some $\boldsymbol{x}_{[0,T]} \in \mathbb{R}^{n\times(T+1)}$.

$\mathcal{E} = \left\{ \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{S} : \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \text{ explains the data } (u_{[0,T-1]}, y_{[0,T-1]}) \right\}$ explaining systems

$\mathcal{E}(n) = \mathcal{E} \cap \mathcal{S}(n)$ explaining systems with n states

$\mathcal{E}(\ell, n) = \mathcal{E} \cap \mathcal{S}(\ell, n)$ explaining systems with lag ℓ and n states

---

True system:

$$\begin{bmatrix} A_{\text{true}} & B_{\text{true}} \\ C_{\text{true}} & D_{\text{true}} \end{bmatrix} \in \mathcal{E}(\ell_{\text{true}}, n_{\text{true}}) \subseteq \mathcal{E}(n_{\text{true}}) \subseteq \mathcal{E}$$

Prior knowledge: $\mathcal{S}_{\text{pk}} \subseteq \mathcal{S}$ with $\begin{bmatrix} A_{\text{true}} & B_{\text{true}} \\ C_{\text{true}} & D_{\text{true}} \end{bmatrix} \in \mathcal{S}_{\text{pk}}$

Upper bounds on the lag and state dimension:

- Recall that

$$\ell_{\text{true}} \leqslant L \quad \text{and} \quad n_{\text{true}} \leqslant N$$

- Define

$$\mathcal{S}_{L,N} := \{ [\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}] \in \mathcal{S}(\ell, n) : \ell \leqslant L \text{ and } n \leqslant N \}$$

- Our prior knowledge is thus:

$$\mathcal{S}_{\text{pk}} = \mathcal{S}_{L,N} \cap \mathcal{M}$$

Definition: The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysId if

- $\mathcal{E} \cap \mathcal{S}_{\text{pk}} = \mathcal{E}(n_{\text{true}}) \cap \mathcal{S}_{\text{pk}}$ (data determine state dimension)
- Any two systems in $\mathcal{E} \cap \mathcal{S}_{\text{pk}}$ are related by a state-space transformation

$\boldsymbol{\ell}_{\mathbf{min}} = \min\{\ell \geqslant 0 : \exists\, n \geqslant 0 \text{ s.t. } \mathcal{E}(\ell, n) \neq \varnothing\}$ minimum lag to explain the data

$\boldsymbol{n}_{\mathbf{min}} = \min\{n \geqslant 0 : \mathcal{E}(n) \neq \varnothing\}$ minimum state dimension to explain the data

---

Theorem: $\mathcal{E}(\ell, n) \neq \varnothing \implies n - \ell \geqslant n_{\min} - \ell_{\min}$ $\implies \ell \leqslant n - n_{\min} + \ell_{\min}$

Observation: $L_{\mathrm{d}} := N - n_{\min} + \ell_{\min}$ data-guided bound on lag

$L_{\mathrm{a}} := \min(L, L_{\mathrm{d}})$ actual upper bound

---

Theorem (Camlibel and Rapisarda, 2024): The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysId if and only if

$$T \geqslant L_{\mathrm{a}} + (L_{\mathrm{a}} + 1)m + n_{\min}$$

and

$$\operatorname{rank} \begin{bmatrix} H_{L_{\mathrm{a}}+1}(u_{[0,T-1]}) \\ H_{L_{\mathrm{a}}+1}(y_{[0,T-1]}) \end{bmatrix} = (L_{\mathrm{a}} + 1)m + n_{\min}.$$

Moreover, if these conditions are satisfied then $\ell_{\mathrm{true}} = \ell_{\min}$ and $n_{\mathrm{true}} = n_{\min}$.

The problem of experiment design

Informativity for system identification

The shortest experiment

Recall: $L_\mathrm{a} := \min(L, N - n_\mathrm{min} + \ell_\mathrm{min})$

Theorem: The data $(u_{[0,T-1]}, y_{[0,T-1]})$ are informative for SysId **if and only if**

$$T \geqslant L_\mathrm{a} + (L_\mathrm{a}+1)m + n_\mathrm{min} \quad \text{and} \quad \operatorname{rank} \begin{bmatrix} H_{L_\mathrm{a}+1}(u_{[0,T-1]}) \\ H_{L_\mathrm{a}+1}(y_{[0,T-1]}) \end{bmatrix} = (L_\mathrm{a}+1)m + n_\mathrm{min}.$$

Moreover, if these conditions are satisfied, then $\ell_\mathrm{true} = \ell_\mathrm{min}$ and $n_\mathrm{true} = n_\mathrm{min}$.

Observation: The **shortest** possible informative data length is

$$\boldsymbol{T} := \boldsymbol{L} + (\boldsymbol{L}+1)m + n_\mathrm{true} \quad \text{where} \quad \boldsymbol{L} := \min(L, N - n_\mathrm{true} + \ell_\mathrm{true})$$

Question: Is it possible to **generate** informative data $(u_{[0,\boldsymbol{T}-1]}, y_{[0,\boldsymbol{T}-1]})$, i.e,

$$\operatorname{rank} \begin{bmatrix} H_{\boldsymbol{L}+1}(u_{[0,\boldsymbol{T}-1]}) \\ H_{\boldsymbol{L}+1}(y_{[0,\boldsymbol{T}-1]}) \end{bmatrix} = (\boldsymbol{L}+1)m + n_\mathrm{true}$$

without knowing ℓ_true and n_true?

For the data $(u_{[0,t-1]}, y_{[0,t-1]})$, define

$$H_k^t = \begin{bmatrix} u(0) & \cdots & u(t-k) \\ \vdots & & \vdots \\ u(k-1) & \cdots & u(t-1) \\ \hdashline y(0) & \cdots & y(t-k) \\ \vdots & & \vdots \\ y(k-1) & \cdots & y(t-1) \end{bmatrix}, \qquad G_k^t = \begin{bmatrix} u(0) & \cdots & u(t-k) \\ \vdots & & \vdots \\ u(k-1) & \cdots & u(t-1) \\ \hdashline y(0) & \cdots & y(t-k) \\ \vdots & & \vdots \\ y(k-2) & \cdots & y(t-2) \end{bmatrix},$$

$$\ell_{\min}^t, \quad n_{\min}^t, \quad \text{and} \quad L_{\mathrm{a}}^t := \min(L, N - n_{\min}^t + \ell_{\min}^t).$$

Main idea: start with $k = 1$ and iterate between the following steps:

- increase the rank of G_k^t until no progress can be made
- increase the depth k by one

Important question: when to stop?

Lemma: We have that

$$\boxed{\operatorname{rank} G_k^t \leqslant m + \operatorname{rank} H_{k-1}^t}$$

Lemma: If

$$\boxed{\operatorname{rank} G_k^t < m + \operatorname{rank} H_{k-1}^t},$$

then there exists an $m-1$ dimensional affine set $\mathcal{A}^t \subseteq \mathbb{R}^m$ such that

$$\operatorname{rank} G_k^{t+1} = \operatorname{rank} G_k^t + 1 \quad \text{whenever} \quad u(t) \notin \mathcal{A}^t.$$

Theorem: Suppose that $(u_{[0,t-1]}, y_{[0,t-1]})$ is such that

- H_k^t has full column rank, and
- $\boxed{\operatorname{rank} G_k^t = m + \operatorname{rank} H_{k-1}^t}$.

Then, $k = L_{\mathrm{a}}^t + 1$ implies that

1. $k = \boldsymbol{L} + 1$,
2. $t = \boldsymbol{T}$, and
3. $(u_{[0,\boldsymbol{T}-1]}, y_{[0,\boldsymbol{T}-1]})$ are **informative for SysId**.

procedure $\textsc{OnlineExperiment}(L, N)$
  choose $u_{[0,m-1]}$ nonsingular
  measure outputs $y_{[0,m-1]}$
  $t \leftarrow m, \;\; k \leftarrow 1$
  while $k \neq L_{\mathrm{a}}^{\boldsymbol{t}} + 1$ **do** ▷ stopping criteria
    $k \leftarrow k + 1$
    if $t = k - 1$ **then**
      choose $u(t)$ arbitrarily ▷ $\boldsymbol{G}_{\boldsymbol{k}}^{\boldsymbol{t}}$ has (full) rank 1
      measure output $y(t)$
      $t \leftarrow t + 1$
    end if
    while $\boxed{\operatorname{rank} \boldsymbol{G}_{\boldsymbol{k}}^{\boldsymbol{t}} < m + \operatorname{rank} \boldsymbol{H}_{\boldsymbol{k-1}}^{\boldsymbol{t}}}$ **do**
      choose $u(t) \notin \boldsymbol{\mathcal{A}}^{\boldsymbol{t}}$ ▷ $\operatorname{rank} \boldsymbol{G}_{\boldsymbol{k}}^{\boldsymbol{t+1}} = \operatorname{rank} \boldsymbol{G}_{\boldsymbol{k}}^{\boldsymbol{t}} + 1$
      measure output $y(t)$
      $t \leftarrow t + 1$
    end while
  end while
  return $(u_{[0,t-1]}, y_{[0,t-1]})$ ▷ $(k, t) = (\boldsymbol{L} + 1, \boldsymbol{T})$ and data are informative
end procedure

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n^1_{\min} = 0$, $\ell^1_{\min} = 0 \implies L^1_{\mathrm{a}} = \min(L, N - n^1_{\min} + \ell^1_{\min}) = 2 \implies k \neq L^1_{\mathrm{a}} + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G^3_2 = \begin{bmatrix} 1 & 0 \\ 0 & u(2) \\ -1 & 0 \\ 2 & 0 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n_{\min}^1 = 0$, $\ell_{\min}^1 = 0 \implies L_{\text{a}}^1 = \min(L, N - n_{\min}^1 + \ell_{\min}^1) = 2 \implies k \neq L_{\text{a}}^1 + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G_2^3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 2 & 0 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n^1_{\text{min}} = 0$, $\ell^1_{\text{min}} = 0 \implies L^1_{\text{a}} = \min(L, N - n^1_{\text{min}} + \ell^1_{\text{min}}) = 2 \implies k \neq L^1_{\text{a}} + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G^4_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & u(3) \\ -1 & 0 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n^1_{\min} = 0$, $\ell^1_{\min} = 0 \implies L^1_{\mathrm{a}} = \min(L, N - n^1_{\min} + \ell^1_{\min}) = 2 \implies k \neq L^1_{\mathrm{a}} + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G^4_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n^1_{\min} = 0$, $\ell^1_{\min} = 0 \implies L^1_{\text{a}} = \min(L, N - n^1_{\min} + \ell^1_{\min}) = 2 \implies k \neq L^1_{\text{a}} + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G^5_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & u(4) \\ -1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 1 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \; B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$u(0) = 1 \implies y(0) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Let $t = 1$ and $k = 1$.

$n^1_{\text{min}} = 0$, $\ell^1_{\text{min}} = 0 \implies L^1_{\text{a}} = \min(L, N - n^1_{\text{min}} + \ell^1_{\text{min}}) = 2 \implies k \neq L^1_{\text{a}} + 1$

Set $k = 2$. Since $t = k - 1$, let $u(1) = 0$ (arbitrary) $\implies y(1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Now **increase rank**:

$$G^5_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2 & 0 & 2 & 1 \end{bmatrix}$$

True system and initial state:

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix},\ B_{\text{true}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ C_{\text{true}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ D_{\text{true}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},\ x(0) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Hence, $n_{\text{true}} = 2$ and $\ell_{\text{true}} = 1$. We take $N = L = 2$.

$$\operatorname{rank} H_1^5 = \operatorname{rank} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 2 & 0 & 2 & 1 & 1 \end{bmatrix} = 3 \implies \boxed{\operatorname{rank} G_2^5 = 1 + \operatorname{rank} H_1^5}$$

$\ell_{\min}^5 = 1$ and $n_{\min}^5 = 2 \implies L_{\text{a}}^5 = \min(2, 2-2+1) = 1 \implies k = L_{\text{a}}^5 + 1$.

Conclusion: The data $(u_{[0,4]}, y_{[0,4]})$ are **informative for SysId**

Reduction in # samples: from $T = 9$ to $T = 7$ to $\boldsymbol{T} = 5$

The shortest experiments for system identification require:

1. **Online design** of the inputs
2. **Online adaptation** of the **depth** of the Hankel matrix

---

Online design using depth-$(L+1)$ Hankel matrix is shortest only if

$$L \leqslant N - n_{\text{true}} + \ell_{\text{true}}$$

Final example: For a system with

$$m = 80, \quad p = 10, \quad \ell_{\text{true}} = 20, \quad n_{\text{true}} = 100,$$

and

$$L = 100, \quad N = 150,$$

- fundamental lemma requires: $T = 20330$
- online design (fixed depth) requires: $T = 8280$
- the shortest experiment requires: $\boldsymbol{T} = 5850$

Thank you!